

AD-A158 996

REALIZATION AND APPROXIMATION OF STATIONARY STOCHASTIC  
PROCESSES(U) MASSACHUSETTS INST OF TECH CAMBRIDGE LAB  
FOR INFORMATION AND D. Y AVNIEL FEB 85 LIDS-TN-1440  
AFOSR-TR-85-0674 DRAG29-84-K-0005

1/1

UNCLASSIFIED

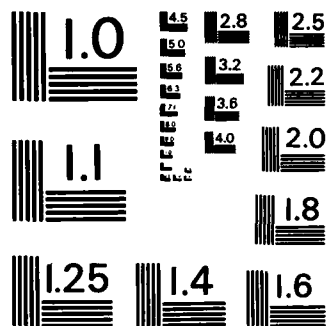
F/G 12/1

NL

END

FILED

DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS - 1963 - A

AFOSR-TR- 85-0674

DTIC  
Apr 85

(2)

FEBRUARY 1985

LIDS-TH-1440

AD-A158 996

Research Supported By:

*Air Force Office of Scientific Research  
Grant AFOSR-82-0135B*

*Army Research Office  
Grant DAAG29-84-K-0005*

REALIZATION AND  
APPROXIMATION OF  
STATIONARY STOCHASTIC  
PROCESSES

Yehuda Avniel

DTIC  
ELECTE  
SEP 12 1985  
S D E

Approved for public release;  
distribution unlimited.

DTIC FILE COPY

Laboratory for Information and Decision Systems  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

85 9 09 080

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A158 996

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>			1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE				
4. PERFORMING ORGANIZATION REPORT NUMBER(S) <b>LIDS-TH-1440</b>			5. MONITORING ORGANIZATION REPORT NUMBER(S) <b>AFOSR-TR- 85-0674</b>	
6a. NAME OF PERFORMING ORGANIZATION Massachusetts Institute of Technology		6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research	
6c. ADDRESS (City, State and ZIP Code) Laboratory for Information and Decision Systems, Cambridge MA 02139			7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION <b>AFOSR</b>		8b. OFFICE SYMBOL (If applicable) <b>NM</b>	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER <b>AFOSR-82-0135</b>	
8c. ADDRESS (City, State and ZIP Code) <b>Bolling AFB DC 20332-6448</b>			10. SOURCE OF FUNDING NOS.	
			PROGRAM ELEMENT NO. <b>61102F</b>	TASK NO. <b>A1</b>
			PROJECT NO. <b>2304</b>	WORK UNIT NO.
11. TITLE (Include Security Classification) <b>REALIZATION AND APPROXIMATION OF STATIONARY STOCHASTIC PROCESSES</b>				
12. PERSONAL AUTHOR(S) <b>Yehuda Avniel</b>				
13a. TYPE OF REPORT <b>Technical</b>		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Yr., Mo., Day) <b>FEB 85</b>
15. PAGE COUNT <b>84</b>				
16. SUPPLEMENTARY NOTATION				
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB. GR.		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)				
<p>To a multivariate stationary stochastic process, we associate a scattering matrix <math>X</math>, which measures the interaction between the past and future of the process. This matrix valued function can be viewed as the generalized phase function associated with the spectral density. It determines the density up to congruency only for a completely non-deterministic sequence.</p> <p>Using the theory of Adamjan-Arov-Krein on extensions of Hankel operators, we establish that the Hankel operator <math>H_S</math> determines the Laurent operator <math>L_S</math> as its unique norm preserving lifting. Employing the Nagy-Foias theory on unitary dilations, or its dual, (CONTINUED)</p>				
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <b>UNCLASSIFIED/UNLIMITED</b> <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>			21. ABSTRACT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>	
22a. NAME OF RESPONSIBLE INDIVIDUAL <b>Dr. Marc Q. Jacobs</b>			22b. TELEPHONE NUMBER (Include Area Code) <b>(202) 767-4940</b>	22c. OFFICE SYMBOL <b>NM</b>

ITEM #19, ABSTRACT, CONTINUED: Lax-Phillips scattering operator model, we develop a realization theory for equivalent classes of stationary sequences with the same density. The minimal equivalence class of Markovian representations is induced by the coprime factorization of the scattering matrix. This presents a unified approach to stochastic and deterministic realization theory, with  $S$  as the analog of the frequency response function.

To obtain reduced order models, we approximate the given sequence with a jointly stationary one of a lower dimensional state space, minimizing the distance between the two sequences. The solution which involves  $H_S$  is non-constructive. We pose a weaker version, leading to a Hankel approximation of  $H_S$ . The algorithm employs the analytic properties of Schmidt pairs for a Hankel operator. An error bound on the normalized difference between the covariance functions of the two sequences is obtained.

February 1985

LIDS-TH-1440

**REALIZATION AND APPROXIMATION  
OF STATIONARY STOCHASTIC PROCESSES**

by

Yehuda Avniel

This report is based on the unaltered thesis of Yehuda Avniel, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in February 1985. This research was conducted at the M.I.T. Laboratory for Information and Decision Systems, with support provided by the Air Force Office of Scientific Research (AFOSR-82-0135B) and the Army Research Office (DAAG29-84-K-0005).

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)  
NOTICE OF APPROVAL FOR PUBLICATION  
This report is approved for publication and distribution as indicated.  
DISTRIBUTION STATEMENT  
MATTHEW J. ...  
Chief, Technical Information Division

Laboratory for Information and Decision Systems  
Massachusetts Institute of Technology  
Cambridge, Massachusetts 02139

REALIZATION AND APPROXIMATION  
OF STATIONARY STOCHASTIC PROCESSES

by

Yehuda Avniel

B.S., Technion - Israel Institute of Technology  
(1976)

M.S., Johns Hopkins University  
(1979)

SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February, 1985

© Massachusetts Institute of Technology, 1985

Signature of Author

*Yehuda Avniel*  
Department of Electrical Engineering  
and Computer Science  
February, 1985

Certified by

*Sanjoy K. Mitter*  
Sanjoy K. Mitter  
Thesis Supervisor

Accepted by

*Arthur C. Smith*  
Arthur C. Smith, Chairman  
Departmental Committee on Graduate Students



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

(A)

REALIZATION AND APPROXIMATION  
OF STATIONARY STOCHASTIC PROCESSES

by

Yehuda Avniel

Submitted to the Department of Electrical Engineering and Computer Science on February 19, 1985, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

To a multivariate stationary stochastic process, <sup>the author</sup> we associate a scattering matrix  $S$ , which measures the interaction between the past and future of the process. This matrix valued function can be viewed as the generalized phase function associated with the spectral density. It determines the density up to congruency only for a completely non-deterministic sequence.

Using the theory of Adamjan-Arov-Krein on extensions of Hankel operators, <sup>this report</sup> ~~we establish~~ <sup>sub S</sup> that the Hankel operator  $H_S$  determines the Laurent operator  $L_S$  as its unique norm preserving lifting. Employing the Nagy-Foias theory on unitary dilations, or its dual, Lax-Phillips scattering operator model, <sup>sub S</sup> we develop a realization theory for equivalent classes of stationary sequences with the same density. <sup>is developed</sup> The minimal equivalence class of Markovian representations is induced by the coprime factorization of the scattering matrix. This presents a unified approach to stochastic and deterministic realization theory, with  $S$  as the analog of the frequency response function.

To obtain reduced order models, <sup>the author</sup> we approximate the given sequence with a jointly stationary one of a lower dimensional state space, minimizing the distance between the two sequences. The solution which involves  $H_S$  is non-constructive. We pose a weaker version, leading to a Hankel approximation of  $H_S$ . The algorithm employs the analytic properties of Schmidt pairs for a Hankel operator. An error bound on the normalized difference between the covariance functions of the two sequences is obtained.

Thesis Supervisor: Dr. Sanjoy K. Mitter  
Title: Professor of Electrical Engineering



**לאבי ז"ל**  
**לאמי שתחיה**

**לא עליך המלאכה לגמור**

**ולא אתה בן חורין להבטל ממנה**

**(פרקי אבות ב'-ט"ז)**

## ACKNOWLEDGEMENTS

I am indebted to my advisor, Sanjoy K. Mitter, for the challenge and stimulation he provided me, the particular interest he paid my work, and his ongoing generous support.

Thanks are due as well to the members of my thesis committee-- Victor Guillemin for his encouragement and support; Bernard C. Levy for his constructive comments; and Gunther Uhlmann for many helpful conversations.

In addition, I wish to thank Donald E. Sarason for bringing to my attention Bloomfield, Jewell, and Hayashi's paper in the Pacific Journal of Mathematics [Reference 9].

I wish to express my gratitude to the Massachusetts Institute of Technology, Department of Electrical Engineering, for awarding me the Vinton Hayes Fellowship in Communications (1980, 1982) and for supporting my research through teaching assistantships.

This research was supported by grants from the Air Force Office of Scientific Research (AFOSR-82-0135B) and the Army Research Office (DAAG29-84-K-0005), administered through the Laboratory for Information and Decision Systems.

My thanks also to Becky Bayerl for her special attention and fine work in typing this thesis.

Finally, I wish to thank Sara, my Beruriah.

## TABLE OF CONTENTS

CHAPTER ONE: . . . . .	6
<u>Introduction</u>	
CHAPTER TWO: . . . . .	14
<u>Notation</u>	
CHAPTER THREE: . . . . .	18
<u>The Scattering Operator Model and the</u>	
<u>Scattering Matrix Associated with a</u>	
<u>Scattering Stochastic Process</u>	
CHAPTER FOUR: . . . . .	28
<u>The Induced Hankel and Toeplitz</u>	
<u>Operators</u>	
CHAPTER FIVE: . . . . .	42
<u>Block Matrix Representations</u>	
CHAPTER SIX: . . . . .	47
<u>Markov Processes and Unitary</u>	
<u>Dilations</u>	
CHAPTER SEVEN: . . . . .	58
<u>Markovian Representations and</u>	
<u>Factorization of the Scattering Matrix</u>	
CHAPTER EIGHT: . . . . .	66
<u>Model Reduction</u>	

## CHAPTER I

### INTRODUCTION

The work of Adamjan-Arov on scattering [2], and their subsequent investigations of Hankel operators and various approximation, interpolation, and extension problems connected with them [1], [3] - [7], has by and large not been used for the study of stationary discrete time stochastic processes (stationary sequences). Our work is a study of the ramifications of their work in the context of stationary sequences. For deterministic linear systems, their work on Hankel norm approximations has recently received considerable attention. In our view, this is due to the fact that the input-output approach gives the Hankel norm approximation a physically appealing interpretation (see Chapter 8). For stationary sequences, however, the probabilistic information is available in the form of the spectral density  $f_{\underline{y}\underline{y}}$  of the  $p$ -dimensional sequence  $\{\underline{y}(n)\}_{-\infty}^{\infty}$  (or equivalently the covariance function) and there is no obvious candidate which plays the role of an input-output map. This is now well known, for example, in the context of stochastic realization theory (cf. also the work of Willems [42]). Looking for an object which plays the role of an input-output map has some conceptual, as well as practical, ramifications. Conceptually, it presents us with a unified approach to deterministic and stochastic realization theory. On a more practical level, it provides us with some justification as to 'what' should be approximated when dealing with the model reduction problem.

In this study, we associate with every centered regular full rank stationary process  $y$  a scattering matrix  $S$  which measures the interaction between the past and future of the process. It may be considered to be a mapping which maps the backward innovations of the process to the forward innovations of the process (see Chapter 3). In the 1-dimensional case, this scattering function is essentially the phase function of the outer factor of the density (and hence can be considered as the phase function associated with the density).

The question as to when the phase function (for a 1-dimensional process) determines the density of the process (up to a positive scalar multiple) was investigated by Levinson-McKean [28]. They proved that this is true iff (if and only if) the process has the completely non-deterministic property, namely, no value in its (strict) future space (that is, the space spanned by the values  $y(1), y(2), \dots$ ) can be linearly predicted without error based on its past space (that is, the space spanned by  $y(0), y(-1), \dots$ ). This property is stronger than regularity in which no future value of the process can be linearly predicted without error based on its past space [9]. For the vector case, not only does the approach in [28] not generalize, but, in addition, it is not clear what is the right notion of a phase (matrix valued) function to associate with a density. That our scattering matrix  $S$  can be viewed as the right generalization can be supported by the fact that  $S$  determines the density up to congruency (that is, up to the form  $K_{yy}^* f(\lambda) K$ ) iff the process is completely non-deterministic (Theorems 4.5, 4.8). Those scattering matrices which correspond to regular full rank completely non-deterministic processes are characterized in Theorem 4.13.

In deterministic systems realization theory, for a causal frequency

response function  $\phi$  we have the general principle [23, p. 24] that the Hankel operator  $H_\phi$  (the composition of the reachability and observability maps) determines (the causal)  $\phi$  up to an additive constant. Assuming  $\phi(0) = [0]$ , the general principle can be rephrased by saying that  $H_\phi$  admits a unique causal lifting to the Laurent operator  $L_\phi$  (input-output map). In the case of stationary sequences the Hankel operator  $H_S$  admits a unique *norm preserving* lifting  $L_S^+$  (Theorem 4.4). This property can be given a variational characterization, namely,  $S$  is the unique solution to <sup>++</sup>

$$\min\{\|F\|_\infty : F \in L_\infty(B(C^P)), c_k(F) = c_k(S), k = -1, -2, \dots\}$$

where  $c_k(F)$  is the  $k$ -th Fourier coefficient of  $F$ .

That the scattering matrix can be viewed as an analog to the frequency response function in systems is further strengthened by the role we show it plays in Markovian representations (realizations) of stationary sequences. A subspace  $X$  in the span  $H_{\underline{y}}$  of  $\underline{y}$  is a Markovian representation for the sequence  $\{\underline{y}(n)\}_{-\infty}^{\infty}$  if the process  $\{U^n X\}_{-\infty}^{\infty}$  ( $U$  being the shift on  $H_{\underline{y}}$ ) has the (weak) Markov property (Definition 6.1), and  $y_1(0), \dots, y_p(0)$  are contained in  $X$ . The dynamical representations are thereby readily obtained (Sections 6,7).

Ruckebusch [37], Lindquist-Picci [23] develop a realization theory for stationary sequences in which every Markovian representation is in

---

<sup>+</sup> In the terminology of [4],  $S$  is the unique minifunction for  $H$ .

<sup>++</sup> Each such  $F$  can be viewed as an extension of  $H_S$  [4]. The parametrization of all those extensions seems to be strongly related to the covariance extension problem.

1-1 correspondence with a factorization of the spectral density.

In our view, having the spectral density as the only probabilistic information available to us implies that two stationary sequences in  $H_Y$  having the same density (a.e.) are indistinguishable (equivalent). Thus, two geometrically (and probabilistically) different Markov processes in  $H_Y$  inducing identical dynamical representations (see Theorems 6.12, 7.1) should be considered equivalent. This equivalence relation between Markov processes is available to us via the characteristic function of Nagy-Foiaş associated with a (completely non-unitary) contraction or its corresponding (completely non-unitary) semigroup of contractions [30]. Indeed, we demonstrate that for a subspace  $X \subset H_Y$  the (full range, regular) process  $\{U^n X\}_{n=-\infty}^{\infty}$  is (weakly) Markov iff the state transition operator  $A$ , which is the compression of  $U$  to  $X$ , has  $U$  as its minimal unitary power dilation in  $H_Y$  (Proposition 6.2), and, moreover,  $A$  is of class  $C_{00}^+$  (Corollary 6.5). Thus  $\{A^n\}_{n \geq 0}$  is a  $C_{00}$  semigroup of contractions. The equivalence relation between Markov processes will be based upon this Markov semigroup. In view of the dual relationship between the Lax-Phillips scattering operator model and dilation theory [2], there is set up a 1-1 correspondence between a Markov process  $\{U^n X\}_{n=-\infty}^{\infty}$  and a Lax-Phillips scattering system (Theorem 6.4). According to a fundamental result of Nagy-Foiaş [30, Th. VI.2.5] the characteristic function of a  $C_{00}$  contraction induces a model space (the functional model), and under the so-called (Nagy-Foiaş) incoming spectral representation, the action of this contraction on the model space is the restricted shift (the model operator). Thus, two contractions having

---

<sup>+</sup> i.e.,  $A^n \rightarrow 0$ ,  $A^{*n} \rightarrow 0$  strongly.

coinciding <sup>+</sup> characteristic functions are unitarily equivalent, and so are their corresponding semigroups. Noting (as was established in [2]) that the characteristic function of A coincides with the scattering matrix  $\Theta_X^{++}$  of the aforementioned scattering system induced by  $\{U^n X\}_{-\infty}^{\infty}$ , we establish in this way an equivalence relation between Markov processes (thus abandoning the 1-1 correspondence with their scattering systems) which is based upon the coincidence of their scattering matrices (characteristic functions), that is, the unitary equivalence between their corresponding Markov semigroups. This enabled us to demonstrate that two stationary sequences are equivalent iff they possess identical dynamical representations induced by equivalent Markov processes which represent them (Theorems 6.12, 7.1). We next show (Theorem 6.10) that there exists a 1-1 correspondence between the equivalent classes of (full range, regular) Markov processes and the purely contractive analytic functions  $\Theta \in H_{\infty}(B(C^P))$  which are inner. Among those inner functions we wish to choose those and only those which are (via the above correspondence) a representation for the stationary sequence  $\{y(n)\}_{-\infty}^{\infty}$  or a sequence equivalent to it. This is accomplished in Theorem 7.2. Thus, the desired family of inner functions are those and only those  $\Theta$ , which factor the scattering matrix  $S: S = \Theta_1^* \Theta_2, \Theta_2 \in H_{\infty}(B(C^P))$ . The left coprime factorization  $S = Q_1^* Q_2$  induces the minimal equivalent class <sup>+++</sup> (Theorem 7.3). This is in direct agreement with deterministic systems realization.

---

<sup>+</sup> Definition 6.8.

<sup>++</sup> This function coincides with the structural function of Lindquist-Picci [29].

<sup>+++</sup> A corresponding result holds for the right factorizations.



Having associated an inner function  $\Theta$  to each Markov process  $\{U^n X\}_{-\infty}^{\infty}$ , there is a natural degree associated with it, namely,  $d(\Theta) = \det \Theta$ .<sup>+</sup> Accordingly we obtain  $d(Q_1) \leq d(\Theta_1)$ , that is, the minimal class is of the weakest degree. Our approach enables us (by inspection) to derive the fact that a Markovian representation is minimal iff it is observable and constructible, and exact observability and constructibility hold iff range  $H_S$  is closed.

The scattering approach was found to be useful in various prediction and interpolation problems of stationary processes. This was demonstrated by Adamjan-Arov in [3]. They consider a general situation in which for two jointly stationary 1-dimensional<sup>++</sup> processes  $\xi, \eta$  we are given the past  $H_{\eta}^{-}(0)$  of  $\eta$ , and the future  $H_{\xi}^{+}(m)$  ( $m \geq 0$ ) of  $\xi$ , and we wish to linearly predict  $\eta(k)$ ,  $k > 0$  based on  $H_{\eta}^{-}(0) V H_{\xi}^{+}(m)$ . This general set-up is applicable to prediction, interpolation, and filtering. By posing the problem in terms of incoming and outgoing data, the corresponding scattering matrix can be computed and the predictor expressed in terms of this scattering matrix. Thus, generally speaking, this approach transforms the projections in the space of values of the processes into the corresponding ones in  $L_2$  in which a solution can be obtained.

Another area in which the scattering matrix appears to have considerable appeal is in model reduction. As the only accessible data is the spectral density  $f_{yy}$ , we would ultimately be concerned with obtaining rational approximations to this density. We formulate (Chapter 8) two problems in model reduction both of which involve the Hankel operator  $H_S$ . In Problem A we wish to approximate the  $p$ -dimensional

---

<sup>+</sup> Its scalar multiple in the terminology of [30].

<sup>++</sup> As they point out, their approach generalizes to the vector case.

process  $\underline{y}$  with a  $p$ -dimensional process  $\hat{\underline{y}}$  of reduced order  $m$ , in such a way that the 'distance' between them

$$\varepsilon = \sup \{ \|\underline{y}_j(n) - \hat{\underline{y}}_j(n)\|_{H_{\underline{y}}} : j = 1, \dots, p, n \in \mathbb{Z} \}$$

is minimized. The motivation for Problem A stems from the fact that for the covariances  $\{C_n\}_{-\infty}^{\infty}$ ,  $\{\hat{C}_n\}_{-\infty}^{\infty}$  of  $\underline{y}$ ,  $\hat{\underline{y}}$  we then obtain

$$\|C_n - \hat{C}_n\| \leq 2\varepsilon\sqrt{p} \left[ \sum_{j=1}^p \sigma_{y_j y_j}^2 \right]^{\frac{1}{2}}.$$

We give a solution (generally non-unique) to the above problem (Theorem 8.2) which, however, can not be obtained in a constructive fashion.

This problem leads naturally to a weaker version of finding an inner function  $Q \in H_{\infty}(B(C^p))$  of degree  $d(Q) \leq m$  such that  $\|H_{QS}\| = \min$ .

(Problem B). This is a Hankel approximation problem in disguise. For the  $p = 1$  dimensional case we obtain a constructive solution in the following fashion: We consider the rank  $m$  Hankel approximant  $H_{\phi_m}^+$  to  $H_S$ . The desired function  $q$  is obtained by

$$H_2^- \ominus qH_2^- = \text{range } H_{\phi_m}$$

and is constructed from the Schmidt pairs corresponding to the  $m$ -th singular value  $s_m(H_S)$ . On the Markov subspace corresponding to  $q$  we project  $y(0)$  and thereby obtain a process  $\hat{\underline{y}}$  whose density  $f_{\hat{\underline{y}}\hat{\underline{y}}}$  is of degree at most  $m$ , and for its moments we obtain the bound (Proposition 8.5)

---

<sup>+</sup> The vector case can not be treated in generality since  $H_{\phi_m}$  is in general not necessarily strictly non-cyclic.

$$\left| \frac{c_n - \hat{c}_n}{c_0} \right| \leq 2s_m(H_S) , \quad n = 0, \pm 1, \dots$$

In closing we mention that various properties of stationary processes were found to be reflected in  $H_S$  (or  $S$ ), such as the strong mixing property, complete non-determinism, and strict non-cyclicity.

## CHAPTER TWO

## NOTATION

$Z$  stands for the set of integers,  $\delta(n)$  for the indicator function of  $\{0\} \subset Z$ ,  $C$  the complex numbers, and for  $a \in C$   $\bar{a}$  denotes the complex conjugate of  $a$ . For a matrix  $A = (a_{ij})_{i,j=1}^p$  we denote by  $A^*$  the Hermitian conjugate of  $A$  :  $A^* = (b_{ij})_{i,j=1}^p$   $b_{ij} = \bar{a}_{ji}$ , and by  $A'$  its transposition. For a family of subsets  $\{M_j\}_j$  of a Hilbert space  $H$ , we denote by  $\bigvee_j M_j$  the smallest closed linear manifold (subspace) that includes each  $M_j$ , and by  $\bigwedge_j M_j$  the greatest subspace contained in each of them (their intersection).  $\bar{M}_j$  denotes the closure of  $M_j$  in  $H$ . For subspaces  $M, N$  of  $H$ ,  $M \ominus N$  denotes the orthogonal complement of  $N$  in  $M$ . For a countable family  $\{M_j\}_j$  of mutually orthogonal subspaces :  $M_i \perp M_j$   $i \neq j$ , we let  $\sum_j \oplus M_j$  be their orthogonal sum.  $P_M$  stands for the orthogonal projection of  $H$  onto the subspace  $M$ . For a bounded linear operator  $A : H_1 \rightarrow H_2$  of Hilbert space  $H_1$  into  $H_2$ , we denote by  $[A]$  the matrix of  $A$  with respect to specified orthonormal bases in  $H_1, H_2$ .  $A|_M$  stands for the restriction of  $A$  to the subspace  $M \subset H_1$ .  $B(H_1, H_2)$  denotes the Banach space of all bounded linear operator from  $H_1$  into  $H_2$  with  $B(H) = B(H, H)$ .

By  $l_2(-\infty, \infty; N)$  we denote the usual Hilbert space of sequences  $\{h_j\}_{j=-\infty}^{\infty}$  with values in (the Hilbert space)  $N$  for which  $\sum_j \|h_j\|_N^2 < \infty$ .  $l_2(0, \infty; N)$ ,  $l_2(-\infty, -1; N)$  are seen naturally as subspaces of  $l_2(-\infty, \infty; N)$ .  $L_2, L_\infty$  will denote respectively the Lebesgue spaces on the circle  $T = \{e^{i\lambda} : \lambda \in [-\pi, \pi]\}$  (with respect to the normalized Lebesgue measure

$\frac{d\lambda}{2\pi}$ ) of square integrable, essentially bounded complex valued functions.

Each function can be viewed as defined as  $[-\pi, \pi]$ . Similarly for the spaces  $L_2(C^p)$ ,  $L_\infty(C^p)$  of functions  $f$  taking values in  $C^p$  for which

$\|f(\cdot)\|_{C^p} \in L_p$ ,  $\|f(\cdot)\|_{C^p} \in L_\infty$  respectively.  $L_\infty(B(C^p))$  is defined analogously for weakly measurable,  $B(C^p)$  valued functions  $f$  for which

$$\text{ess. sup} \{ \|f(e^{i\lambda})\|_{B(C^p)} : \lambda \in [-\pi, \pi] \} < \infty.$$

$H_2^\pm$  are the subspaces of  $L_2$  defined by

$$H_2^+ = \left\{ f \in L_2 : \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\lambda}) e^{-in\lambda} d\lambda = 0, \quad n = -1, -2, \dots \right\}$$

$$H_2^- = \left\{ f \in L_2 : \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\lambda}) e^{-in\lambda} d\lambda = 0, \quad n = 0, 1, 2, \dots \right\},$$

and we have the orthogonal decomposition  $L_2 = H_2^+ \oplus H_2^-$ . Each  $f \in H_2^+$  having a Fourier series

$$f(e^{i\lambda}) \sim \sum_0^\infty a_n e^{in\lambda}$$

generates the function

$$g(z) = \sum_0^\infty a_n z^n$$

belonging to the Hardy class  $H_2$  of functions  $g(z)$  holomorphic in  $|z| < 1$  and such that

$$\|g\|_{H_2} = \sup_{0 < r < 1} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(re^{i\lambda})|^2 d\lambda \right]^{\frac{1}{2}} < \infty$$

Moreover, the (a.e. existing) radial limit  $g(e^{i\lambda})$  of  $g(z)$  equals  $f(e^{i\lambda})$  a.e. and  $\|f\|_{L_2} = \|g\|_{H_2}$ . The function  $g(z)$  is seen as the analytic extension of  $f \in H_2^+$  to the unit disc  $|z| < 1$  and is denoted by  $f(z)$ . We identify  $H_2^+$  with  $H_2$  and denote them commonly by  $H_2$ . Using the conjugation with respect to the unit circle ( $z \rightarrow \frac{1}{\bar{z}}$ ), by the reflection principle, for  $f \in H_2 \subset L_2$  the function  $\bar{f}$  defined by  $\bar{f}(e^{i\lambda}) = \overline{f(e^{i\lambda})}$  has an analytic extension to  $|z| > 1 : \overline{f(\frac{1}{\bar{z}})}$ , which we again denote by  $\bar{f}$ . The space  $\bar{H}_2 = \{f \in L_2 : \bar{f} \in H_2\}$  is the space of functions  $f \in L_2$  having an analytic extension to the exterior of the disc  $f(z)$  and we have

$$\|f\|_{L_2} = \sup_{\rho > 1} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\rho e^{i\lambda})|^2 d\lambda \right]^{\frac{1}{2}}.$$

$f \in \bar{H}_2$  are called conjugate analytic.

Analogously for the Banach space  $L_\infty$  we have the subspaces  $H_\infty = H_\infty^+ \subset L_\infty$  of functions  $f \in L_\infty$  having an analytic extension  $f(z)$  to  $|z| < 1$  with

$$\|f\|_{L_\infty} = \sup_{|z| < 1} |f(z)| = \|f\|_{H_\infty}.$$

Similarly, for the Hilbert space  $L_2(C^P)$  we have the subspaces  $H_2^+(C^P) = H_2(C^P)$ ,  $H_2^-(C^P)$  with the orthogonal decomposition  $L_2(C^P) = H_2(C^P) \oplus H_2^-(C^P)$ . In  $L_\infty(B(C^P))$ , again  $H_\infty(B(C^P))$  is defined as the subspace of functions in  $L_\infty(B(C^P))$  whose negatively indexed (matrix valued) Fourier coefficients vanish. For  $\Theta \in H_\infty(B(C^P))$  the function  $\Theta^*$  defined by  $\Theta^*(e^{i\lambda}) = [\Theta(e^{i\lambda})]^*$  is identified with its analytic extension  $\Theta^*(\frac{1}{\bar{z}}) = [\Theta(\frac{1}{\bar{z}})]^*$  to  $|z| > 1$ .

A function  $f \in H_\infty$  is called *inner* if  $|f(e^{i\lambda})| = 1$  a.e. . Similarly for  $\Theta \in H_\infty(B(C^p))$  if  $\Theta(e^{i\lambda})$  is unitary a.e. .  $f \in H_2$  is called *outer* if  $V\{\chi^n f\}_{n \geq 0} = H_2$  where  $\chi$  denotes the function on  $T$  defined by  $\chi(e^{i\lambda}) = e^{i\lambda}$ . For  $\phi \in L_\infty(B(C^p))$  the Toeplitz operator  $T_\phi : H_2(C^p) \rightarrow H_2(C^p)$  whose matrix is block Toeplitz with respect to the standard basis  $\{e^{ik\lambda} e_1, e^{ik\lambda} e_2, \dots, e^{ik\lambda} e_p\}_{k \geq 0}$ ,  $\{e_1, e_2, \dots, e_p\}$  being the standard basis in  $C^p$ , is defined by  $T_\phi f = \pi_+(\phi f)$  where  $\pi_+$  is the Riesz projection of  $L_2(C^p)$  onto  $H_2(C^p)$ .  $H_\phi$  will denote the Hankel operator [with block Hankel matrix with respect to the standard bases in  $H_2(C^p)$ ,  $H_2^-(C^p)$ ]  $H_\phi : H_2(C^p) \rightarrow H_2^-(C^p)$  defined by  $H_\phi f = \pi_-(\phi f)$ ,  $\pi_-$  being the Riesz projection of  $L_2(C^p)$  onto  $H_2^-(C^p)$ . The convention we employ regarding a Hankel operator as acting from  $H_2(C^p)$  into  $H_2^-(C^p)$  is not in accordance with the one employed in systems theory, in which we act on  $H_2^-(C^p)$  into  $H_2(C^p)$  :  $H_\phi f = \pi_+(\phi f)$ . It, however, conforms with the one employed by Adamjan-Arov-Krein and enables us to use their results without modifications, as well as to refer to them.

## CHAPTER THREE

THE SCATTERING OPERATOR MODEL AND THE SCATTERING  
MATRIX ASSOCIATED WITH A STATIONARY STOCHASTIC PROCESS

Let  $H$  be a complex separable Hilbert space and let  $U$  be a unitary operator on  $H$ . A subspace  $D_+$  is said to be *outgoing* for  $(U, H)$  if it satisfies

$$(3.1)_+ \quad \begin{aligned} (i) \quad & UD_+ \subset D_+ \\ (ii) \quad & \bigcap_{n=-\infty}^{\infty} U^n D_+ = \{0\} \\ (iii) \quad & \bigvee_{n=-\infty}^{\infty} U^n D_+ = H \end{aligned}$$

A subspace  $D_-$  for which

$$(3.1)_- \quad \begin{aligned} (i) \quad & U^* D_- \subset D_- \\ (ii) \quad & \bigcap_{n=-\infty}^{\infty} U^n D_- = \{0\} \\ (iii) \quad & \bigvee_{n=-\infty}^{\infty} U^n D_- = H \end{aligned}$$

is said to be *incoming* for  $(U, H)$ .

**3.1 Definition.** A quadruple  $(U, H, D_+, D_-)$  satisfying (3.1) is said to be a *scattering system*.

We shall be interested in a scattering system arising in the following way.

Let  $(\Omega, \mathcal{A}, P)$  be a fixed probability space and let

$$\{\underline{y}(n) : n \in \mathbb{Z}\} \quad , \quad \underline{y}(n) = \begin{pmatrix} y_1(n) \\ y_2(n) \\ \vdots \\ y_p(n) \end{pmatrix}$$



be a centered stationary process with  $y_j(n) \in \mathcal{L}_2(\Omega, \mathcal{A}, P)$   $j = 1, \dots, p$ .

Let  $f_{\underline{y}\underline{y}}(\lambda) = (f_{kj}(\lambda))_{k,j=1}^p$ ,  $\lambda \in [-\pi, \pi]$  be its spectral density satisfying

$$(3.2) \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \det f_{\underline{y}\underline{y}}(\lambda) d\lambda > -\infty,$$

i.e., the process is *regular* and of *maximal rank*.

Let

$$H = H_{\underline{y}} = \bigvee_{n \in \mathbb{Z}} \{y_1(n), y_2(n), \dots, y_p(n)\} \subset \mathcal{L}_2(\Omega, \mathcal{A}, P),$$

be the space spanned by the process and let  $U$  be the unitary shift operator on  $H$  associated with the  $\underline{y}$  process [36, p. 14]:

$$U y_j(n) = y_j(n+1) \quad j = 1, \dots, p, \quad n \in \mathbb{Z}.$$

We consider the *past* and *future* of  $\{\underline{y}(n)\}_{-\infty}^{\infty}$  defined by

$$D_- \doteq H_{\underline{y}}^-(0) = \bigvee_{k \leq 0} \{y_1(k), \dots, y_p(k)\}, \quad D_+ \doteq H_{\underline{y}}^+(0) = \bigvee_{k \geq 0} \{y_1(k), \dots, y_p(k)\}.$$

By (3.1) it follows [36, Th. II.6.1]

$$\bigwedge_{-\infty}^{\infty} U^n D_- = \{0\} = \bigwedge_{-\infty}^{\infty} U^n D_+.$$

We readily obtain that  $(U, H, D_+, D_-)$  is a scattering system.

### 3.2 Theorem (Translation Representation Theorem [27, Th. II.1.1]).

Let  $(U, H, D_+)$  be outgoing. Then there exists a Hilbert space  $N_+$  and a unitary map  $r_+$  of  $H$  onto  $l_2(-\infty, \infty; N_+)$  such that

$$(3.3) \quad \begin{aligned} (i) \quad r_+[D_+] &= l_2(0, \infty; N_+) , \\ (ii) \quad U_+ &= r_+ U r_+^{-1} \end{aligned}$$

is the right shift operator on  $l_2(-\infty, \infty; N_+)$ . This representation is unique up to automorphisms of  $N_+$ .

Proof (cf. [2, p. 77]). By (3.1)<sub>+</sub>-ii the operator  $U|_{D_+}$  is an isometry having no unitary part. By Wold's decomposition theorem [30, Th.I.1.1] we may write uniquely

$$(3.4) \quad D_+ = \sum_{n=0}^{\infty} \oplus U^n N_+ \quad N_+ = D_+ \ominus U D_+ .$$

Since for any  $m > 0$

$$\begin{aligned} U^{-m} D_+ &= U^{-m} [(D_+ \ominus U^m D_+) \oplus U^m D_+] = U^{-m} \left[ \left( \sum_{k=0}^{m-1} \oplus U^k N_+ \right) \oplus U^m D_+ \right] = \\ &= \left( \sum_{k=-1}^{-m} \oplus U^k N_+ \right) \oplus D_+ , \end{aligned}$$

we obtain by (3.1 -iii)<sub>+</sub> that

$$(3.5) \quad H = \sum_{-\infty}^{\infty} \oplus U^n N_+ .$$

It follows that for arbitrary  $h \in H$

$$h = \sum_{-\infty}^{\infty} \oplus U^n P_{N_+} U^{-n} h , \quad \|h\|_H^2 = \sum_{-\infty}^{\infty} \|P_{N_+} U^{-n} h\|_H^2 .$$

Hence the map

$$r_+ : H \rightarrow l_2(-\infty, \infty; N_+)$$

defined by

$$(3.6) \quad r_+ h = \{P_{N_+} U^{-n} h\}_{n=-\infty}^{\infty}$$

is isometric. Since for  $\{h_n\}_{n=-\infty}^{\infty} \in l_2(-\infty, \infty; N_+)$ ,  $h = \sum_{n=-\infty}^{\infty} U^n h_n \in H$ , the map  $r_+$  is onto and thus unitary. By (3.4) we obtain (i). From (3.6)

$$r_+ U h = P_{N_+} \{U^{-(n-1)} h\}_{n=-\infty}^{\infty} = U_+(r_+ h) .$$

and (ii) follows. By (3.5)  $U$  is a bilateral shift of multiplicity equal to  $\dim N_+$  and the uniqueness follows.  $\blacktriangleleft$

**3.3 Definition.** The representation  $(U_+, l_2(0, \infty; N_+), l_2(-\infty, \infty; N_+))$  is called an *outgoing translation representation*.

For  $(U, H, D_-)$  incoming we similarly obtain

$$(3.7) \quad D_- = \sum_{n=-\infty}^0 \oplus U^n N_- \quad N_- = D_- \ominus U^* D_- ,$$

and

$$(3.8) \quad H = \sum_{n=-\infty}^{\infty} \oplus U^n N_- .$$

For the corresponding map  $r_-$  of  $H$  onto  $l_2(-\infty, \infty; N_-)$  we define

$$(3.9) \quad r_- h = \{P_{N_-} U^{-(n+1)} h\}_{n=-\infty}^{\infty}$$

Thus

$$(i) \quad r_-[D_-] = l_2(-\infty, -1; N_-) \quad ,$$

$$(ii) \quad U_- = r_- U r_-^{-1}$$

is the right shift on  $l_2(-\infty, \infty; N_-)$ . The representation  $(U_-, l_2(-\infty, -1; N_-), l_2(-\infty, \infty; N_-))$  is called an *incoming translation representation*.

Now let  $(U, H, D_+, D_-)$  be the scattering system associated with the regular maximal rank  $y$  process. The subspace  $N_- = D_- \ominus U^* D_-$  ( $N_+ = D_+ \ominus U D_+$ ) is the *forward (backward) innovation subspace* at  $n = 0$ . Since for a scattering system  $(U, H, D_+, D_-)$  we have

$$\dim N_- = \text{multiplicity } U = \dim N_+ \quad ,$$

we can arrange the maps  $r_{\pm}$  to be onto  $l_2(-\infty, \infty; C^P)$ .

3.4 Definition ([2], [27]). The operator

$$S = r_- r_+^{-1} : l_2(-\infty, \infty; C^P) \rightarrow l_2(-\infty, \infty; C^P)$$

is called the *abstract scattering operator*.

Clearly  $S$  is unitary. Denoting by  $V$  the right shift on  $l_2(-\infty, \infty; C^P)$ , we readily obtain by the translation representation theorem.

$$(3.10) \quad SV = r_- r_+^{-1} V = r_- U r_+^{-1} = VS \quad .$$

Let  $F : l_2(-\infty, \infty; C^P) \rightarrow L_2(C^P)$  be the Fourier transform operator. The unitary operator

$$FSF^{-1} : L_2(C^P) \rightarrow L_2(C^P)$$

thus commutes by (3.10) with  $L_\chi$ , the operator of multiplication by  $\chi$ . It follows [11] that  $FSF^{-1}$  is a Laurent operator  $L_S$ ,  $S \in L_\infty(B(C^P))$  such that

$$S(e^{i\lambda})$$

a.e. is a unitary map on  $C^P$ .

3.5 Definition ([2],[27]).  $S$  is called the *scattering matrix*.

It is clear from the translation representation theorem that  $S$  is determined to within right and left multiplication by unitary transformations on  $C^P$  (i.e., to within coincidence, see Definition 6.8).

We next compute the scattering matrix  $S$  for the  $y$  process. Let  $\{v_1(0), \dots, v_p(0)\}$  be an orthonormal basis for  $N_-$ . Let  $v_j(n) = U^n v_j(0)$  and define

$$\underline{v}(n) = \begin{pmatrix} v_1(n) \\ \vdots \\ v_p(n) \end{pmatrix} \quad n \in \mathbb{Z}.$$

By (3.8), the process  $\{\underline{v}(n)\}_{n=-\infty}^{\infty}$  is a (centered) white noise process with covariance  $R_{\underline{v}\underline{v}}(n) = \delta(n)I_{C^P}$ , constituting the forward innovation process for the  $y$  process. It is determined up to a choice of basis in  $N_-$ . By (3.7), we may write

$$(3.11) \quad \underline{y}(0) = \sum_{k=-\infty}^{\infty} A(k) \underline{v}(k) \quad A(k) = (\alpha_{ij}(k))_{i,j=1}^P \quad A(k) = [0], k > 0.$$

(Wold's representation cf. [36, p. 56]). It follows from (3.9)

$$r_{y_j}(0) = \left\{ \sum_{m=1}^p \alpha_{jm}(k+1) v_m(0) \right\}_{k=-\infty}^{\infty}$$

Identifying  $N_-$  with  $C^p$  we readily obtain the representation

$$(3.12) \quad r_{-}y_j(0) = \left\{ \begin{pmatrix} \alpha_{jl}^{(k+1)} \\ \vdots \\ \alpha_{jp}^{(k+1)} \end{pmatrix} \right\}_{k=-\infty}^{\infty}$$

Consider the function

$$\Lambda(z) = \sum_{k=-\infty}^{\infty} A'(k) z^k$$

Since

$$\sum_{k=-\infty}^{\infty} \sum_{i,j=1}^p |\alpha_{ij}(k)|^2 \leq \sum_{j=1}^p \|y_j(0)\|_H^2 < \infty,$$

$\Lambda(z)$  is analytic in  $|z| > 1$ . For  $\Lambda(z)$  we have [36, p. 57]

$$(3.13) \quad \frac{1}{2\pi} \Lambda^*(z) \Lambda(z) = f_{yy}(\lambda)$$

By the incoming properties

$$H_2^-(C^p) = \bigvee_{n < 0} \{ e^{in\lambda} \Lambda(e^{i\lambda}) \underline{a} : \underline{a} \in C^p \},$$

i.e.,  $\Lambda$  is conjugate outer [21, p. 121]. From (3.12)

$$(Fr_{-}y_1(0), \dots, Fr_{-}y_p(0)) = \bar{\chi} \Lambda.$$

Since the translates (in  $H_{\underline{y}}$ ) of  $y_1(0), \dots, y_p(0)$  and their linear combinations are dense in  $H_{\underline{y}}$ ,  $Fr_{-}$  is determined by the above expression.

We now consider the outgoing representation. Let  $\{\epsilon_1(0), \dots, \epsilon_p(0)\}$  be an orthonormal basis in  $N_+$ . We similarly obtain

$$\underline{y}(0) = \sum_{-\infty}^{\infty} B(k) \underline{\epsilon}(k) \quad B(k) = (\beta_{ij}(k))_{i,j=1}^p, \quad B(k) = [0] \quad , \quad k < 0.$$

This representation constitutes the representation of  $y(0)$  in terms of the backward innovation process  $\{\underline{\epsilon}(n)\}_{-\infty}^{\infty}$ ,  $\underline{\epsilon}(n) = \begin{pmatrix} \epsilon_1(n) \\ \vdots \\ \epsilon_p(n) \end{pmatrix}$ .

We define

$$\Gamma(z) = \sum_0^{\infty} B'(k) z^k$$

which is analytic in  $|z| < 1$ . In a similar fashion to (3.13), we obtain by direct computation

$$(3.15) \quad \frac{1}{2\pi} \Gamma^*(z) \Gamma(z) = f_{\underline{y}\underline{y}}(\lambda) \quad z = e^{i\lambda} \quad ,$$

with  $\Gamma$  being outer. Also

$$(3.16) \quad (Fr_{+}y_1(0), \dots, Fr_{+}y_p(0)) = \Gamma \quad .$$

Combining (3.14) with (3.16), we obtain

$$S\Gamma = \bar{\chi}\Lambda$$

and thus

$$S = \bar{\chi}\Lambda\Gamma^{-1} \quad .$$

Using (3.13) and (3.15), one easily verifies that  $S(e^{i\lambda})$  is unitary

a.e.  $\lambda \in [-\pi, \pi]$ . We thus obtained

3.6 Theorem. For a regular maximal rank process  $\{y(n)\}$  we have

$$(3.17) \quad S = \bar{\chi} \Lambda \Gamma^{-1},$$

where  $S$  is determined up to left and right multiplication by constant unitary matrices.

For the case  $p = 1$  we have

3.7 Corollary. For a regular process  $\{y(n)\}_{-\infty}^{\infty}$

$$(3.18) \quad s = \frac{\bar{\chi}}{\chi} \frac{\bar{\Gamma}}{\Gamma},$$

and  $s$  is determined up to multiplication by a constant of unit modulus.

Proof. The outer function  $\bar{\Lambda}$  satisfies  $|\bar{\Lambda}| = |\Gamma|$  on  $T$  and thus  $\bar{\Lambda} = \gamma \Gamma$  a.e. where  $\gamma$  is a constant of unit modulus. ◀

3.8 Remark. The scattering matrix  $S$  was defined by an outer and conjugate outer factors of the density  $f_{yy}$ . Since those are determined up to left multiplication by a constant unitary matrix, we may wish to make a canonical choice (which amounts to choosing specific orthonormal bases in  $N_+, N_-$ ) in the following fashion: For  $\Gamma(0)$  we consider its polar decomposition  $\Gamma(0) = KP$  ( $K$  unitary,  $P > 0$ ) and define  $\Gamma_1(z) = K^{-1}\Gamma(z)$ . For  $\Gamma_1$  we have  $\Gamma_1(0) > 0$  and (3.15) holds. This  $\Gamma_1$  is unique. Similarly for  $\Lambda$ . In this way, the density  $f_{yy}$  will have a unique  $S$  associated with it. From the viewpoint of seeing  $S$  as the phase function associated with  $f_{yy}$ , this may be appealing. Note, however, that  $S$  measures the



interaction between the past and future of the process  $\underline{y}$  (see discussion following theorem 4.5) and uniqueness in the reverse direction (from  $S$  to  $f_{\underline{y}\underline{y}}$ ) is not possible.

## CHAPTER FOUR

### THE INDUCED HANKEL AND TOEPLITZ OPERATORS

We call the unitary maps  $F_- = Fr_-$ ,  $F_+ = Fr_+$  - the *incoming* and *outgoing spectral representations*, respectively.

4.1 Proposition. We have

- (i)  $F_-[D_-] = H_2^-(C^P)$  ,
- (ii)  $F_+[D_+] = SH_2(C^P)$  ,
- (iii)  $F_-(Uh) = F_-h \quad h \in H_y$  .

Proof. (i) and (iii) follow from the properties of the incoming translation representation. By the definition of S we obtain

$$F_-[D_+] = (Fr_-r_+^{-1}F^*)Fr_+[D_+] = SH_2(C^P) .$$

4.2 Lemma. The operator

$$P_-P_+ : D_+ \rightarrow D_- , \quad P_{\pm} = P_{D_{\pm}}$$

is unitarily equivalent to the Hankel operator  $H_S$  , and the operator

$$P_{D_-}^+ P_+ : D_+ \rightarrow D_-^+$$

is unitarily equivalent to the Toeplitz operator  $T_S$  .

Proof. For  $h \in D_+$  and  $f = F_+ h$  we have

$$F_-(P_-P_+)F_+^*f = F_-P_-h = F_-P_-F_-^*F_-F_+^*f = (F_-P_-F_-^*)Sf .$$

Since

$$F_-P_-F_-^* = \pi_-$$

we obtain

$$F_-(P_-P_+)F_+^*f = H_S f \quad f \in H_2 ,$$

and the following diagram commutes

$$\begin{array}{ccc}
 D_+ & \xrightarrow{P_-P_+} & D_- \\
 F_+ \downarrow & & \downarrow F_- \\
 H_2(C^P) & \xrightarrow{H_S} & H_2^-(C^P)
 \end{array}$$

The unitarity of  $F_-, F_+$  implies that  $P_-P_+$  is unitary equivalent to  $H_S$ .

This proves the first part. Now since

$$F_-(I_H - P_-)F_+^*f = F_-F_+^*f - F_-P_-F_+^*f = Sf - \pi_-Sf = \pi_+Sf = T_S f ,$$

the following diagram commutes

$$\begin{array}{ccc}
 D_+ & \xrightarrow{P_{D_-^\perp}} & D_-^\perp \\
 F_+ \downarrow & & \downarrow F_- \\
 H_2(C^P) & \xrightarrow{T_S} & H_2(C^P)
 \end{array}$$

which proves the second part.

4.3 Lemma. For  $H_S$ ,  $T_S$  we have the identity

$$H_S^* H_S + T_S^* T_S = I_{H_2(C^P)}.$$

Proof. Clearly,

$$H_S^* g = \pi_+(S^* g) \quad g \in H_2^-(C^P).$$

It follows that for  $f \in H_2(C^P)$ :

$$\begin{aligned}
 H_S^* H_S f &= \pi_+ S^* [\pi_-(Sf)] = \pi_+ S^* [(I_{L_2(C^P)} - \pi_+) Sf] = \\
 &= \pi_+ S^* Sf - \pi_+ S^* \pi_+ Sf = f - T_S^* T_S f.
 \end{aligned}$$

By a theorem of Nehari [31] (a vector generalization of which was obtained by Sarason [41]), for a bounded Hankel operator  $H_\phi$ ,  $\phi \in L_\infty$  there exists a function  $\phi_\mu \in L_\infty$  such that  $H_\phi = H_{\phi_\mu}$  and

$$(4.1) \quad \|H_\phi\| = \|\phi_\mu\|_\infty$$

Since always  $\|H_\phi\| \leq \|\phi_\mu\|_\infty$  because of (4.1),  $\phi_\mu$  is called a *minifunction* for  $H_\phi$  [5, p. 6]. The question regarding the uniqueness of the minifunction is of particular interest to us. In system realization theory for a frequency response function  $\phi \in H_\infty$  (same reasoning holds for the vector case), a central role is played by the Hankel operator  $H_\phi$  :  $H_2^- \rightarrow H_2$ ,  $H_\phi f = \pi_+(\phi f)$ . Now observe that if  $\phi_1, \phi_2 \in H_\infty$  are such that  $H_{\phi_1} = H_{\phi_2}$  then  $H_{\phi_1 - \phi_2} \equiv 0$ , and the positive Fourier coefficients of  $\phi_1 - \phi_2$  vanish. Thus  $\phi_1 - \phi_2 = \text{const.}$  and  $H_\phi$  determines (the analytic)  $\phi$  up to an additive constant (recall in this regard that the composition of the reachability and observability maps determine the frequency response function up to an additive constant). However, in general, (4.1) does not hold and the uniqueness of  $\phi$  inducing  $H_\phi$  is guaranteed by analyticity (causality). In view of the central role played by  $S$  in realization of stationary sequences (Chapter 7), the following theorem is of significance:

**4.4 Theorem.** The Hankel operator  $H_S$  determines  $S$  uniquely. Indeed,  $S$  is its unique minifunction.

Proof. From Lemma 4.3 we note that  $f \in \text{Ker } T_S$  iff  $f$  is an eigenvector of  $H_S^* H_S$  corresponding to the eigenvalue  $\|H_S\| = 1$ . Since  $S = \bar{\chi} \Lambda \Gamma^{-1}$  every column of  $\Gamma$  belongs to this kernel. Thus, the projection of the above eigenspace on the first coordinate in  $l_2(0, \infty; C^p)$  spans  $\Gamma(0)$ . Now observe that for  $\Gamma(0)$  we have, because of its outer property in  $H_2(B(C^p))$ , (see e.g. [36, p. 76])

$$\log \frac{|\det \Gamma(0)|}{(2\pi)^{p/2}} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \det f_{\underline{y}\underline{y}}(\lambda) d\lambda > -\infty,$$

so that  $\Gamma(0)$  is of full rank. We conclude that the aforementioned projection is onto the first coordinate space. According to a result of Adamjan-Arov-Krein [4, Corollary 3.1] for a Hankel operator  $H_\phi$  to have a unique minifunction, it is sufficient that the projection of the eigenspace of  $H_\phi^* H_\phi$  corresponding to  $\|H_\phi\|$  on the first coordinate space be onto. The result follows. ◀

There is an alternative way to rephrase Theorem 4.4. If we consider the Laurent operator  $L_{\phi_\mu}$  of multiplication by  $\phi_\mu$  on  $L_2$ , then (4.1) becomes

$$\|H_\phi\| = \|L_{\phi_\mu}\|$$

Since

$$H_\phi = P_{H_2} L_{\phi_\mu} |_{H_2}$$

one considers  $L_{\phi_\mu}$  as norm preserving lifting of  $H_\phi$ . Thus, in system theory, the uniqueness of the lifting is guaranteed by causality, the lifting being in general not norm preserving, while for stationary processes, the uniqueness is guaranteed by the lifting being norm preserving.

Viewing a linear time invariant system from the input-output point of view makes the frequency response function the sole accessible object containing all pertinent information about the system. As to the information contained in the scattering matrix, we have the following:

**4.5 Theorem.** The scattering matrix  $S$  determines the density  $f_{\underline{y}\underline{y}}(\lambda)$  up to the form

$$(4.2) \quad K^*_{\underline{f}_{\underline{y}\underline{y}}}(\lambda)K$$

where  $K$  is a constant  $p \times p$  non-singular matrix, iff

$$(4.3) \quad \dim \text{Ker } T_S = p \quad .$$

Proof. First note that for any representation of  $S$

$$S = \bar{Y}YX^{-1}$$

with the columns of  $X$  in  $H_2(C^p)$  and those of  $\bar{X}Y$  in  $H_2^-(C^p)$ , the columns of  $X$  belong to  $\text{Ker } T_S$ . Moreover (on  $T$ )

$$Y^*Y = (SX)^*SX = X^*X \quad .$$

Assume (4.3) holds. It thus follows that

$$(4.4) \quad X(e^{i\lambda}) = \Gamma(e^{i\lambda})K$$

where  $K$  is a  $p \times p$  full rank constant matrix. Thus,

$$\frac{1}{2\pi} X^*(z)X(z) = \frac{1}{2\pi} K^* \Gamma^*(z) \Gamma(z) K = K^*_{\underline{f}_{\underline{y}\underline{y}}}(\cdot) K \quad z = e^{i\lambda} \quad ,$$

proving the 'if' part.

Now assume (4.3) not to hold, i.e.,  $\dim \text{Ker } T_S > p$ . We can thus find a  $p \times p$  matrix  $X(e^{i\lambda})$  of full rank a.e.  $\lambda$  such that the columns of  $X$  belong to  $\text{Ker } T_S$  and (4.4) does not hold. If we define

$$Y = \bar{X}SX$$

then the columns of  $\bar{X}Y$  are in  $H_2^-(C^P)$  and  $S = \bar{X}YX^{-1}$  with  $Y^*Y = X^*X$ .

The result follows.  $\blacktriangleleft$

That the scattering matrix determines  $f_{\underline{y}\underline{y}}$  up to the form (4.2) is a natural consequence of the scattering framework. Indeed, for an arbitrary non-singular  $K$ , the process  $\underline{\xi}(n) = K^* \underline{y}(n)$  whose density equals

$$f_{\underline{\xi}\underline{\xi}} = K^* f_{\underline{y}\underline{y}} K$$

induces the same scattering system  $(U, H, D_+, D_-)$  as the  $\underline{y}$  process.

We next characterize condition (4.3) on a process level.

4.6 Proposition. We have

$$F_+^*[\text{Ker } T_S] = H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(0)$$

Proof. Let  $0 \neq f \in \text{Ker } T_S$ . From Lemma 4.3 it follows

$$H_S^* H_S f = f$$

i.e.,

$$\|H_S f\| = \|f\|$$

By Lemma 4.2, we obtain for  $\underline{\xi} = F_+^* f \in H_{\underline{y}}^+(0)$

$$\|P_- \underline{\xi}\| = \|\underline{\xi}\| ,$$

and  $\underline{\xi} \in H_{\underline{y}}^-(0)$ . Thus

$$F_+^*[\text{Ker } T_S] \subset H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(0) .$$



Now let  $\underline{\xi} \in H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(0)$ . It follows from Lemma 4.2

$$H_S(F_+\xi) = F_-\xi.$$

Let  $f = F_+\xi \in H_2(C^p)$ . We obtain

$$\|H_S f\| = \|F_-\xi\| = \|\xi\| = \|F_+\xi\| = \|f\|$$

and

$$H_S^* H_S f = f.$$

Thus  $f \in \text{Ker } T_S$  which implies

$$F_+[H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(0)] \subset \text{Ker } T_S.$$

The result follows. ◀

By the unitarity of  $F_+$

$$\dim \text{Ker } T_S = \dim H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(0),$$

and since  $\underline{y}$  is regular and of full rank we readily conclude

$$\dim H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(0) = p \text{ iff } \dim H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(1) = 0.$$

**4.7 Definition [9].** The process  $\underline{y}$  is said to be *completely non-deterministic* if

$$H_{\underline{y}}^-(0) \wedge H_{\underline{y}}^+(1) = \{0\}$$

As is well-known (see e.g. [36, p. 73]) for a regular maximal rank process  $\{\underline{y}(n)\}_{-\infty}^{\infty}$  no  $y_j(k)$   $k \geq 1$ ,  $j = 1, \dots, p$  can be predicted without error based on the past  $H_{\underline{y}}^-(0)$ . Being completely non-deterministic is more restrictive; indeed, no value in  $H_{\underline{y}}^+(1)$  can be predicted without error based on  $H_{\underline{y}}^-(0)$ .

Summarizing, we can restate theorem 4.5 in the following way.

**4.8 Theorem.** The scattering matrix  $S$  determines  $f_{\underline{y}\underline{y}}$  up to the form (4.2) iff  $\underline{y}$  is completely non-deterministic.

**4.9 Remark.** It is of interest to observe that since for a completely non-deterministic process, the eigenvectors of  $H_S^* H_S$  corresponding to  $\|H_S\|$  are only the columns of  $\Gamma$ , the projection of this eigenspace on the first coordinate is not only onto, but also 1-1. In [4, Sec. 2] it is shown that for any Hankel operator  $H : H_2(C^p) \rightarrow H_2^-(C^p)$  satisfying this condition, its unique minifunction is of the form

$$\rho S \qquad \rho = \|H\|$$

Thus up to a constant multiple  $\rho > 0$  all minifunctions of such Hankel operators are in 1-1 correspondence with regular, full rank, completely non-deterministic processes.

**4.10 Example.** Let  $\{y(n)\}_{-\infty}^{\infty}$  have rational density

$$f_{\underline{y}\underline{y}}(\lambda) = \frac{|P(z)|^2}{|Q(z)|^2} \qquad z = e^{i\lambda},$$

where the polynomials  $P, Q$  have no zeros in  $|z| < 1$  and are relatively prime. Since  $f_{\underline{y}\underline{y}} \in L_1$ , the polynomial  $Q$  has its zeros in  $|z| > 1$ . Write

$$P = P_1 P_2$$

where  $P_1$  of degree  $k$  has its zeros on  $T$  and  $P_2$  in  $|z| > 1$ . For

$$P_1(z) = \prod_{j=1}^k (z - \alpha_j) \quad , \quad \{\alpha_j\}_{j=1}^k \subset T \text{ we have}$$

$$\frac{\overline{P_1(e^{i\lambda})}}{P_1(e^{i\lambda})} = e^{-ik\lambda} (-1)^k \prod_{j=1}^k \bar{\alpha}_j \quad .$$

Thus

$$(4.5) \quad s = \gamma \bar{\chi}^{k+1} \frac{\bar{\psi}_e}{\psi_e} \quad \psi_e = \frac{P_2}{Q}, \quad \gamma = (-1)^k \prod_{j=1}^k \bar{\alpha}_j \quad ,$$

where  $|\gamma| = 1$  and  $\psi_e$  is outer. In [5] Adamjan-Arov-Krein show that (4.5) is the general form of unimodular minifunctions and that in this case  $k+1$  is the dimension of the eigenspace corresponding to the singular value  $1 = \|H_S\|$  [5, Th. 2.2]. From Lemma 4.3 it readily follows that this dimension equals  $\dim \text{Ker } T_S$ . Thus <sup>+</sup>

$$\dim H_y^-(0) \Delta H_y^+(0) = k+1 \quad .$$

We conclude that a regular process with rational density is completely non-deterministic iff it has no zeros on  $T$ .

4.11 Example. Let  $\{y(n)\}_{n=-\infty}^{\infty}$  have density

$$0 < m \leq f_{yy}(\lambda) \leq M < \infty \quad \text{a.e.} \quad .$$

It readily follows that the outer factor  $\Gamma \in H_{\infty}$  and, moreover,  $1/\Gamma \in H_{\infty}$ .

---

<sup>+</sup> This reproduces a result of Bloomfield-Jewel-Hayashi [2, Th.11].

Now let  $g \in \text{Ker } T$ , for arbitrary  $f \in H_2$

$$0 = (T_S g, f) = (\overline{\lambda} \overline{\Gamma} g, f) = (\frac{g}{\Gamma}, \lambda f) .$$

Since  $\Gamma$  is outer and  $g/\Gamma \in H_2$ , we conclude

$$g/\Gamma \in H_2 \ominus \lambda H_2 = \{0\} .$$

Thus,

$$\text{Ker } T_S = \{\lambda \Gamma : \lambda \in \mathbb{C}\}$$

and  $\dim \text{Ker } T_S = 1$ . The process  $y$  is, therefore, completely non-deterministic.

We now formulate a converse to Theorem 4.5. The question can be posed as follows: Under what conditions is a function  $S \in L_\infty(B(C^p))$  a scattering matrix of some full rank,  $p$  dimensional, completely non-deterministic process. On an abstract level we first observe that any  $S \in L_\infty(B(C^p))$  which is unitary valued a.e. on  $T$  is the scattering matrix of the canonical scattering system [2]

$$U = L_\chi, H = L_2(C^p), D_+ = S H_2(C^p), D_- = H_2^-(C^p).$$

The above questions amounts to characterizing all scattering systems  $(U, H, D_+, D_-)$  for which there exists a set  $\{\xi_1, \dots, \xi_p\}$  of linearly independent vectors such that

$$H = \overline{\text{span } \{U^n \xi_j : j=1, \dots, p, n=0, \pm 1, \dots\}}$$

$$(4.6) \quad D_+ = \overline{\text{span } \{U^n \xi_j : j=1, \dots, p, n \geq 0\}}$$

$$D_- = \overline{\text{span } \{U^n \xi_j : j=1, \dots, p, n \leq 0\}}$$

and such that any other linearly independent set satisfying (4.6) is of cardinality  $p$ . The corresponding process will be

$$\{\underline{\xi}(n)\}_{n=-\infty}^{\infty} \text{ where } \underline{\xi}(0) = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}, \quad \underline{\xi}(n) = \begin{pmatrix} U^n \xi_1 \\ \vdots \\ U^n \xi_p \end{pmatrix}, \text{ and the spectral density is}$$

$$\text{obtained by } f_{\underline{\xi}\underline{\xi}}(\lambda) = \left( \frac{d(E_\lambda \xi_i, \xi_j)_H}{d\lambda} \right)_{i,j=1, \dots, p}, \quad \{E_\lambda : \lambda \in [-\pi, \pi]\}$$

being the resolution of the identity for  $U$ . The answer is given in the following.

**4.13 Theorem.** Let  $S \in L_\infty(B(C^p))$  be such that

$$(i) \quad S(e^{i\lambda}) \text{ is a.e. } \lambda \text{ a unitary map on } C^p,$$

$$(ii) \quad \dim \text{Ker } T_S = p.$$

Then there exists a  $p$ -dimensional full rank completely non-deterministic process  $y$  whose scattering matrix is  $S$ .

**Proof.** Let  $\Gamma_1, \Gamma_2, \dots, \Gamma_p$  span the kernel of  $T_S$  and define

$$\Gamma = [\Gamma_1 | \Gamma_2 | \dots | \Gamma_p]$$

Let

$$\Lambda = S\Gamma$$

Since  $\Lambda_j = \pi_-(S\Gamma_j) + \pi_+(S\Gamma_j) = \pi_-(S\Gamma_j)$ ,  $j=1, \dots, p$ , the columns of  $\Lambda = [\Lambda_1 | \Lambda_2 | \dots | \Lambda_p]$  are in  $H_2^-(C^p)$  and by (i)

$$\Lambda^*(z)\Lambda(z) = \Gamma^*(z)\Gamma(z) \quad z = e^{i\lambda}.$$

If we define

$$f_{yy}(\lambda) = \frac{1}{2\pi} \Gamma^*(e^{i\lambda}) \Gamma(e^{i\lambda})$$

the theorem follows provided we show that  $\Gamma$  is outer and  $\chi\Lambda$  conjugate outer. Let  $U = L_\chi$  and define:

$$\hat{D}_- = \bigvee_{n \leq -1} \{\chi^n \Lambda_1, \dots, \chi^n \Lambda_p\} \subset H_2^-(C^p), \quad \hat{D}_+ = \bigvee_{n \geq 0} \{\chi^n S\Gamma_1, \dots, \chi^n S\Gamma_p\} \subset SH_2(C^p).$$

Let

$$(4.7) \quad \hat{H} = \left( \bigvee_{n \in \mathbb{Z}} U^n \hat{D}_- \right) \vee \left( \bigvee_{n \in \mathbb{Z}} U^n \hat{D}_+ \right)$$

It is easily verified that (3.1) $_{\pm}$  - (i), (ii) holds for  $(U, D_{\pm})$ . In [2] Adamjan-Arov show [2, Th. 2.5] that a quadruple  $(U, H, D_+, D_-)$  satisfying (3.1) $_{\pm}$  - (i), (ii) and (4.7) has a scattering matrix  $\hat{S}$  which is unitary valued a.e. on  $T$  iff

$$\bigvee_{n \in \mathbb{Z}} U^n \hat{D}_+ = H = \bigvee_{n \in \mathbb{Z}} U^n \hat{D}_-$$

and, moreover, from their generalized functional model [2, Th. 2.1] we need have

$$D_- = H_2^-(C^p), \quad D_+ = \hat{S}H_2(C^p).$$

A straightforward computation (mimicking the one in section 3) gives

$$\hat{S} = S$$

and the result follows. ◀

4.14 Remark. We wish to comment here on a conceptual point. Note, that the fact that  $S$  (or  $H_S$ ) are used subsequently in various model reduction problems, and realization, is in no way contingent upon  $S$  determining  $f_{yy}$ . Indeed, in all those cases we use explicitly  $f_{yy}$  in that we are able to identify  $F_{-y}(0) = \bar{\chi}_L$  or  $F_{+y}(0) = \Gamma$ .

# CHAPTER FIVE BLOCK MATRIX REPRESENTATIONS

In order to economize notation, we will denote the matrix

$$\left( (\xi_i, \eta_j)_H \right)_{i,j=1,\dots,p} \text{ by } \langle \underline{\xi}, \underline{\eta} \rangle, \quad \underline{\xi} = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}, \quad \underline{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_p \end{pmatrix}. \quad \text{Also, as before,}$$

$$U_{\underline{\xi}} = \begin{pmatrix} U_{\xi_1} \\ \vdots \\ U_{\xi_p} \end{pmatrix}.$$

Recall that  $\{v(n)\}_{-\infty}^{\infty}$ ,  $\{\underline{e}(n)\}_{-\infty}^{\infty}$  are the  $^+$  forward and backward innovation processes for  $\{y(n)\}_{-\infty}^{\infty}$ . Let

$$\tau_{ij} = \langle \underline{e}(j), \underline{v}(i) \rangle \quad i, j \in \mathbb{Z}.$$

By stationarity

$$\tau_{ij} = \langle U^j \underline{e}(0), U^i \underline{v}(0) \rangle = \langle \underline{e}(0), U^{i-j} \underline{v}(0) \rangle = \tau_{i-j, 0}.$$

Denote

$$\gamma_n = \tau'_{n+1, 0} \quad n \in \mathbb{Z}.$$

We form the two-way infinite Toeplitz block matrix

---

$^+$  They are determined to within a choice of orthonormal bases in  $N_-, N_+$  which we fix. See also Remark 3.8.



$$\Delta = \begin{bmatrix} \ddots & & & & \\ & \gamma_{-2} & \gamma_{-1} & \gamma_0 & \gamma_1 & \gamma_2 \\ & \ddots & & & & \\ & & \gamma_{-2} & \gamma_{-1} & \gamma_0 & \gamma_1 & \gamma_2 \\ & & \ddots & & & & \\ & & & \gamma_{-2} & \gamma_{-1} & \gamma_0 & \gamma_1 & \gamma_2 \\ & & & \ddots & & & & \ddots \end{bmatrix}$$

From the definition of the scattering operator via the innovations representations we obtain for the Laurent operator  $L_S$  (with respect to the standard basis)

$$[L_S] = \Delta.$$

As to the Toeplitz operator  $T_S$ , observe that with respect to the orthonormal bases

$$\{v_1(n), v_2(n), \dots, v_p(n)\}_{n \geq 1}, \{\varepsilon_1(n), \varepsilon_2(n), \dots, \varepsilon_p(n)\}_{n \geq 0}$$

in  $D_-^+$ ,  $D_+$  the operator  $P_{D_-^+}|_{D_+}$  has the representation

$$\Sigma = \begin{bmatrix} \langle \underline{\varepsilon}(0), \underline{v}(1) \rangle & \langle \underline{\varepsilon}(1), \underline{v}(1) \rangle & \langle \underline{\varepsilon}(2), \underline{v}(1) \rangle & \dots \\ \langle \underline{\varepsilon}(0), \underline{v}(2) \rangle & \langle \underline{\varepsilon}(1), \underline{v}(2) \rangle & \dots & \dots \\ \langle \underline{\varepsilon}(0), \underline{v}(3) \rangle & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & \dots \\ \gamma_1 & \gamma_0 & \dots & \dots \\ \gamma_2 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

and from lemma 4.2

$$[T_S] = \Sigma$$

Similarly, for  $H_S$  we note that with respect to the orthonormal bases

$$\{v_1(n), v_2(n), \dots, v_p(n)\}_{n \leq 0}, \{\varepsilon_1(k), \varepsilon_2(k), \dots, \varepsilon_p(k)\}_{k \geq 0}$$

in  $D_-$ ,  $D_+$ , the operator  $P_-P_+$  has the matrix representation

$$\Xi = \begin{bmatrix} \langle \underline{\varepsilon}(0), \underline{v}(0) \rangle' & \langle \underline{\varepsilon}(1), \underline{v}(0) \rangle' & \langle \underline{\varepsilon}(2), \underline{v}(0) \rangle' \\ \langle \underline{\varepsilon}(0), \underline{v}(-1) \rangle' & \langle \underline{\varepsilon}(1), \underline{v}(-1) \rangle' & \dots \\ \langle \underline{\varepsilon}(0), \underline{v}(-2) \rangle' & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} = \begin{bmatrix} \gamma_{-1} & \gamma_{-2} & \gamma_{-3} \dots \\ \gamma_{-2} & \gamma_{-3} & \dots \\ \gamma_{-3} & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

so that by Lemma 4.2

$$[H_S] = \Xi.$$

Let

$$[L_S] = \left[ \begin{array}{c|c} [L_S]_{11} & [L_S]_{12} \\ \hline [L_S]_{21} & [L_S]_{22} \end{array} \right]$$

be the partition of  $[L_S]$  corresponding to the orthogonal decomposition  $L_2(C^p) = H_2(C^p) \oplus H_2^-(C^p)$ . By our previous considerations,

$$[L_S]_{21} = [H_S], \quad [L_S]_{11} = [T_S].$$

$[L_S]_{12}$  is the matrix representation of the Hankel operator mapping  $H_2^-(C^p) \rightarrow H_2(C^p)$  taking  $f \mapsto \pi_+(Sf)$ . Similarly, the Toeplitz matrix  $[L_S]_{22}$  which represents the map  $H_2^-(C^p) \rightarrow H_2^-(C^p)$  takes  $f \mapsto \pi_-(Sf)$ . Combining with Theorem 4.4, we obtain

5.1 Theorem. The scattering operator  $S$  has the two way infinite Toeplitz block matrix representation

$$[L_S] = \begin{bmatrix} [T_S] & * \\ [H_S] & ** \end{bmatrix}$$

where  $[H_S]$  determines  $[L_S]$  uniquely.

We next obtain a representation for  $[H_S]$  in terms of the moments

$$C_n = \langle \underline{y}(n), \underline{y}(0) \rangle = \int_{-\pi}^{\pi} \underline{f}_{\underline{y}\underline{y}}(\lambda) e^{in\lambda} d\lambda \in B(C^p), \quad n \in \mathbb{Z}.$$

Using the representation of  $\underline{y}(n)$  in terms of the backward and forward innovations (Chapter 3), we obtain

$$\begin{aligned} (5.1) \quad C_n &= \langle \underline{y}(n), \underline{y}(0) \rangle = \left\langle \sum_{k=0}^{\infty} B(k) \underline{\varepsilon}(k+n), \sum_{m=0}^{\infty} A(-m) \underline{v}(-m) \right\rangle \\ &= \sum_{k=n}^{\infty} \sum_{m=0}^{\infty} B(k-n) \gamma'_{-(m+k+1)} A^*(-m). \end{aligned}$$

Direct computation gives

$$(5.2) \quad \sum_{k=0}^{\infty} B(k) B^*(k+n) = C_n = \sum_{m=n}^{\infty} A(-m) A^*(-m+n) \quad n=0, 1, \dots$$

Form the triangular block matrices

$$B = \begin{bmatrix} B'(0) & & & 0 \\ B'(1) & B'(0) & & \\ B'(2) & B'(1) & B'(0) & \\ \dots & \dots & \dots & \dots \end{bmatrix}, \quad A = \begin{bmatrix} A'(0) & & & 0 \\ A'(-1) & A'(0) & & \\ A'(-2) & A'(-1) & A'(0) & \\ \dots & \dots & \dots & \dots \end{bmatrix}.$$

From (5.2)

$$B^*B = \begin{bmatrix} C_0 & C_1 & C_2 & \dots \\ C_{-1} & C_0 & C_1 & \dots \\ C_{-2} & C_{-1} & C_0 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}, \quad A^*A = \begin{bmatrix} C_0 & C_{-1} & C_{-2} & \dots \\ C_1 & C_0 & C_{-1} & \dots \\ C_2 & C_1 & C_0 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

In accordance with (3.13), (3.15) we call  $B$ ,  $A$ , the outer, conjugate outer factorizations of the Toeplitz form induced by the moments

$\{C_n\}_0^\infty$  ( $C_{-n} = C_n^*$ ). Combining with (5.1) gives

$$A^*[H_S]B = \begin{bmatrix} C'_0 & C'_1 & C'_2 & \dots \\ C'_1 & C'_2 & \dots & \dots \\ C'_2 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}.$$

## CHAPTER SIX

### MARKOV PROCESSES AND UNITARY DILATIONS

In a Hilbert space setting, a centered stationary process  $\{\underline{x}(n)\}_{-\infty}^{\infty}$   $\underline{x}(n) = \begin{pmatrix} x_1(n) \\ \vdots \\ x_m(n) \end{pmatrix}$  is said to be Markov if for all  $n \geq s$

$$(6.1) \quad P_{H_{\underline{x}}^-(s)} \underline{x} = P_{X(s)} \underline{x} \quad \underline{x} \in H_{\underline{x}}^+(s),$$

where  $X(s) = \text{span}\{x_j(s) : j=1, \dots, m\}$ . In our setting, all stationary processes will be generated by the shift  $U$  (on  $H_{\underline{y}}$ ) associated with the  $\underline{y}$  process. Thus, for a stationary process  $\{\underline{x}(n)\}_{-\infty}^{\infty}$  (in  $H_{\underline{y}}$ ) we will have  $\underline{x}(n) = U^n \underline{x}(0)$ . It readily follows from (6.1) that one can define the notion of a Markov subspace with respect to  $U$  in the following (see [23], [37]).

**6.1. Definition.** A subspace  $X \subset H_{\underline{y}}$  is said to be Markov with respect to  $U$  if for all  $n \geq s$ ,  $\underline{x} \in X$

$$(6.2) \quad P_{\bigvee_{-\infty}^s U^m X} U^n \underline{x} = P_{U^s X} U^n \underline{x}.$$

Thus  $X$  is a Markov subspace with respect to  $U$  iff the process  $\{U^n X\}_{-\infty}^{\infty}$  has the (weak) Markov property. In what follows a Markov process  $\{U^n X\}_{-\infty}^{\infty}$  will invariably arise in this fashion.

We shall be interested in Markov subspaces (with respect to  $U$ )<sup>+</sup>

$X \subset H_{\underline{y}}$  for which

$$(6.3) \quad \{y_1(0), \dots, y_p(0)\} \subset X.$$

In this case  $X$  satisfies

$$H_{\underline{y}} = \bigvee_{-\infty}^{\infty} U^n X,$$

and we say that  $X$  (or  $\{U^n X\}_{-\infty}^{\infty}$ ) is of *full range*. There is a direct relationship between Markov processes of full range and unitary dilations. Recall [30] that a unitary operator  $U$  on a Hilbert space  $H$  is said to be the minimal unitary (power) dilation of a contraction  $A$  on  $X \subset H$  if

$$A^n = P_X U^n|X \quad n \geq 0$$

and  $H = \bigvee_{-\infty}^{\infty} U^n X$  (*minimality*).

**6.2 Proposition.**  $X \subset H_{\underline{y}}$  is a Markov subspace of full range iff  $U$  (on  $H_{\underline{y}}$ ) is the minimal unitary (power) dilation of

$$A = P_X U|X : X \rightarrow X.$$

Proof. From (6.2), we obtain for  $x, x' \in X$  and  $m, n \geq 0$

$$(U^{-m}x, U^n x') = (U^{-m}x, P_X U^n x').$$

---

<sup>+</sup> We shall subsequently omit the phrase "with respect to  $U$ ".

Thus, denoting  $A(n) = P_X U^n|X$ , we obtain

$$\begin{aligned} (x, A(m+n)x') &= (x, U^{m+n}x') = (U^{-m}x, U^n x') = (U^{-m}x, P_X U^n x') = \\ &= (x, P_X U^m P_X U^n x') = (x, A(m)A(n)x') \quad . \end{aligned}$$

We conclude that  $A(m+n) = A(m)A(n)$  and

$$A(n) = A^n(1) = A^n.$$

Since  $X$  is of full range, i.e.,  $H_{\underline{y}} = \bigvee_{-\infty}^{\infty} U^n X$ , we conclude that  $U$  in  $H_{\underline{y}}$  is the minimal unitary dilation of  $A$  (in  $X$ ). This proves the 'only if' part. The 'if' part follows by reversing the argument. ◀

The semigroup  $\{A^n\}_{n \geq 0} = \{P_X U^n|X\}_{n \geq 0}$  will be called the *Markov semigroup*.

Much in the same way as for a regular process we make the following:

**6.3 Definition.** A Markov process  $\{U^n X\}_{-\infty}^{\infty} \subset H_{\underline{y}}$  is said to be *regular* if

$$\bigwedge_{n \leq 0} \bigvee_{k \leq n} U^n X = \{0\} = \bigwedge_{n \geq 0} \bigvee_{k \geq n} U^k X$$

We shall occasionally also refer to (the Markov subspace)  $X$  as being *regular*.

The notion of regularity for a Markov process is intimately related to the asymptotic stability of the static operator  $A$  (see Corollary 6.5 and Theorem 6.11).

The correspondence between Markov processes and scattering systems is established in the following:

**6.4 Theorem.** Let  $X \subset H_{\underline{y}}$  be a regular Markov subspace of full range.

Then  $H_{\underline{y}}$  decomposes and, moreover, uniquely into the orthogonal sum

$$(6.4) \quad H_{\underline{y}} = D_{-} \oplus X \oplus D_{+}$$

where  $(U, H_{\underline{y}}, D_{+}, D_{-})$  is a scattering system.

Proof. Define

$$(6.5) \quad \begin{aligned} D_{+} &= \left( \bigvee_{n \geq 0} U^n X \right) \ominus X \\ D_{-} &= \left( \bigvee_{n \leq 0} U^n X \right) \ominus X \end{aligned} .$$

We first show

$$(6.6) \quad U D_{+} \subset D_{+}$$

Note that

$$D_{+} = \bigvee_{n \geq 0} (I_{H_{\underline{y}}} - P_X) U^n X .$$

For arbitrary  $x, x' \in X$  and  $n \geq 0$

$$(U(I_{H_{\underline{y}}} - P_X) U^n x, x') = (U^{n+1} x, x') - (U P_X U^n x, x') =$$

$$(A^{n+1} x, x') - (U A^n x, x') = (A^{n+1} x, x') - (A A^n x, x') = 0 .$$

It follows that

$$U D_{+} \perp X ,$$

and since  $U D_{+} \subset \bigvee_{n \geq 0} U^n X$  we conclude (6.6). Similarly, we obtain  $U^* D_{-} \subset D_{-}$ .



To prove that  $D_- \perp D_+$  we note that it suffices to prove

$$D_+ \perp \bigvee_{m \leq 0} U^m X.$$

Since for arbitrary  $x, x' \in X$ , and  $n, m \geq 0$

$$\begin{aligned} ((I_{H_Y} - P_X)U^n x, U^{-m} x') &= (U^n x, U^{-m} x') - (P_X U^n x, U^{-m} x') = \\ &= (A^{m+n} x, x') - (A^m A^n x, x') = 0 \end{aligned}$$

the conclusion follows. To prove that  $\bigwedge_{-\infty}^{\infty} U^n D_{\pm} = \{0\}$  observe that

$$\bigvee_{-\infty}^{\infty} U^n X = H_Y \Leftrightarrow \bigwedge_{-\infty}^{\infty} U^n X^{\perp} = \{0\} \Leftrightarrow \bigwedge_{-\infty}^{\infty} U^n (D_- \oplus D_+) = \{0\} \Leftrightarrow$$

$$(\bigwedge_{-\infty}^{\infty} U^n D_-) \oplus (\bigwedge_{-\infty}^{\infty} U^n D_+) = \{0\}.$$

By regularity, we obtain that on the space  $K_+ = \bigvee_{n \geq 0} U^n X$  the operator  $U_+ = U|_{K_+}$  is an isometry having no unitary part (Wold's decomposition [30, Th. I.1.1]). Moreover, from the structure of the space of the minimal isometric dilation (of  $A$ ) [30, section II.2], we obtain

$$K_+ = \sum_{n=0}^{\infty} \oplus U^{n+1} N_- \quad N_- = D_- \ominus U^* D_-.$$

Since  $X$  is of full range, we have  $H_Y = \sum_{-\infty}^{\infty} \oplus U^n N_-$  and

$$(6.7)_- \quad \bigvee_{-\infty}^{\infty} U^n D_- = H_Y.$$

In a similar way,

$$(6.7)_+ \quad \bar{V} \sum_{-\infty}^{\infty} U^n D_+ = H_{\underline{y}}.$$

We therefore established that  $(U, H_{\underline{y}}, D_+, D_-)$  is a scattering system. To prove uniqueness, note that if

$$H_{\underline{y}} = D'_- \oplus X \oplus D'_+$$

is another decomposition, then  $UX + UD'_- \supset D'_-$  and, thus  $UX \subset X \oplus D'_+$ . It follows  $D_+ \subset X \oplus D'_+$ , since  $D_+ \perp X$  we obtain

$$D_+ \subset D'_+.$$

Similarly,

$$D_- \subset D'_-.$$

However,  $D_- \oplus D_+ = D'_- \oplus D'_+$  and uniqueness follows.

**6.5 Corollary.** A full range Markov subspace is regular iff <sup>+</sup>

$$A^n \rightarrow 0, \quad A^{*n} \rightarrow 0 \quad (n \rightarrow \infty)$$

strongly.

**Proof.** From the proof of the last theorem it follows that  $X$  is regular iff (6.7) holds. Combining with [30, Th.II.1.2] gives the desired result.

**6.6 Definition.** A scattering system  $(U, H, D_+, D_-)$  for which

---

<sup>+</sup> For a finite dimensional  $X$  both convergences are equivalent.

$$D_- \perp D_+$$

is called a *Lax-Phillips* (L-P) scattering system.

Let  $\{U^n X\}_{n=-\infty}^{\infty}$  be an arbitrary regular Markov process of full range, and  $(U, H_Y, D_+, D_-)_X$  its associated L-P scattering system. Let  $\Theta_X(e^{i\lambda})$  be the corresponding scattering matrix. For the corresponding incoming spectral representation  $F_X^-$ , we obtain from Proposition 4.1

$$F_X^- [D_-] = H_2^-(C^P) \quad , \quad F_X^- [D_+] = \Theta_X H_2(C^P) \quad .$$

Since  $D_+ \perp D_-$

$$\Theta_X \in H_{\infty}(B(C^P)) \quad .$$

To each regular full range Markov process there is thus associated an inner function  $\Theta_X$ , which is the scattering matrix of the corresponding L-P system  $(U, H_Y, D_+, D_-)_X$ . According to Proposition 6.2 and Theorem 6.4, characterizing those inner functions amounts to characterizing all those L-P systems  $(U, H, D_+, D_-)$  for which  $U$  on  $H$  is the minimal unitary dilation of the contraction  $A = P_X U|_X$ ,  $X = H \ominus (D_- \oplus D_+)$ . By [27, Th. III.1.1],  $U$  on  $H$  is a dilation, and combining with [2, Th.3.3], we obtain

**6.7 Proposition.** The scattering matrices  $\Theta$  associated with regular full range Markov processes are those and only those inner functions  $\Theta \in H_{\infty}(B(C^P))$  for which

$$(6.8) \quad \|\Theta(0)\| < 1 \quad .$$

An inner function  $\Theta \in H_\infty(B(C^p))$  satisfying (6.8) is called *purely contractive* [30, p. 188]. For  $p = 1$ , this amounts to being non-trivial.

Recall that the scattering matrix was defined up to left and right multiplication by constant unitary matrices on  $C^p$ . This follows from the arbitrariness in choice of orthonormal bases in the forward and backward innovation subspaces, and cannot be avoided. We make the following

6.8 Definition [30, p. 132].  $\Theta \in H_\infty(B(C^p))$ , and  $\Theta_1 \in H_\infty(B(C^p))$  are said to *coincide* if for unitary maps  $\tau_1, \tau_2$  on  $C^p$

$$\tau_2 \Theta_1(z) \tau_1 = \Theta(z) \quad |z| < 1.$$

The equivalence relation between Markov processes is obtained by

6.9 Definition. We say that the Markov processes  $\{U^n X\}_{-\infty}^\infty, \{U^n X_1\}_{-\infty}^\infty$  are *equivalent* if  $\Theta_X, \Theta_{X_1}$  coincide.

It readily follows from [30, Th. VI. 2.3] that the corresponding Markov semigroups are unitarily equivalent.

While the correspondence

$$X \longleftrightarrow (U, H_Y, D_+, D_-)_X$$

is 1-1, the same does not hold for  $\Theta_X$ , two distinct Markov processes may have coinciding scattering matrices. However, as we shall see in the next section, if we see (as we do in our context) two processes  $Y, Y_1$ , as equivalent if they have the same density  $f_{YY}$  then the above equivalence relation is the right one.

Combining with Proposition 6.7, we obtain

6.10 Theorem. All equivalence classes of full range regular Markov processes are parametrized by the inner functions  $\Theta \in H_\infty(B(C^p))$  such

that  $\|\theta(0)\| < 1$ .

We next obtain a dynamical representation for a Markov process.

**6.11 Theorem.** Let  $X$  be a Markov subspace. Let  $\{x_1(0), \dots, x_k(0)\}$  ( $k \leq \infty$ ) be a complete orthonormal basis in  $X$ . Then for  $\underline{x}(n) = U^n \underline{x}(0)$ ,  $\underline{x}(0) = \begin{pmatrix} x_1(0) \\ \vdots \\ x_k(0) \end{pmatrix}$  we have

$$(6.9) \quad \underline{x}(n+1) = [A]\underline{x}(n) + [B]\underline{w}(n) \quad n \geq 0$$

where

(i)  $\{\underline{w}(n)\}_{n=-\infty}^{\infty}$  is a normalized <sup>+</sup> white noise vector process of dimension  $p$ .

(ii)  $\{\underline{w}(n)\}_{n \geq 0}$  is orthogonal to  $\underline{x}(0)$ .

The above representation determines  $[A]$  (the matrix representation of the state operator  $A$ ) and  $[B]$  up to unitary equivalence.

Proof. Let  $(U, H_y, D_+, D_-)_X$  be the corresponding scattering system. Write  $D_+ = \sum_0^{\infty} \oplus U^n N_+$ ,  $N_+$  the backward innovation subspace at  $t = 0$  for  $D_+$ ,  $\dim N_+ = p$ . Let the entries of  $\underline{w}(0) = \begin{pmatrix} w_1(0) \\ \vdots \\ w_p(0) \end{pmatrix}$  be an orthonormal basis in  $N_+$  with  $\underline{w}(n) = U^n \underline{w}(0)$ . Since  $U^n N_+ \perp N_+$  ( $n \neq 0$ )  $\{\underline{w}(n)\}_{n=-\infty}^{\infty}$  is a normalized white noise process, and  $X \perp D_+$  implies that  $\{\underline{w}(n)\}_{n \geq 0}$  is orthogonal to  $\underline{x}(0)$ . From (6.5)

$$UX \subset X \oplus N_+$$

---

<sup>+</sup> Its covariance matrix equals  $\delta(n)I_{Cp}$ .

and thus

$$(6.10) \quad Ux_j(0) = P_X Ux_j(0) + P_{N_+} Ux_j(0) = Ax_j(0) + P_{N_+} Ux_j(0) .$$

Let

$$\alpha_{ij} = (Ax_i(0), x_j(0))_{H_y}, \quad b_{ij} = (Ux_i(0), w_j(0))_{H_y},$$

and define

$$[A] = (\alpha_{ij}), \quad [B] = (b_{ij}) .$$

From (6.10)

$$\underline{x}(1) = [A]\underline{x}(0) + [B]\underline{w}(0) ,$$

applying U to both sides gives

$$\underline{x}(n+1) = [A]\underline{x}(n) + [B]\underline{w}(n) .$$

Since the above representation is unique up to a choice of orthonormal bases in  $X$ ,  $N_+$  the theorem follows. ◀

The notion of equivalence between Markov processes is naturally manifested in their dynamical representations:

**6.12 Theorem.** If  $\{U^n X\}_{-\infty}^{\infty}$ ,  $\{U^n X_1\}_{-\infty}^{\infty}$  are equivalent Markov processes (with  $X$  as in Theorem 6.11), then  $X_1$  admits representation (6.9) for some  $\underline{x}_1$ ,  $\underline{w}_1$  playing the role of  $\underline{x}$ ,  $\underline{w}$  in Theorem 6.11.

**Proof.** Let  $(U, H_y, D_+, D_-)_X$ ,  $(U, H_y, D'_+, D'_- )_{X_1}$  be the corresponding

scattering systems, and  $F_X^-$ ,  $F_{X_1}^-$  the incoming spectral representations.

It readily follows that for the unitary operator

$$(6.11) \quad L = (F_{X_1}^-)^{-1} F_X^-$$

we have:

$$(i) \quad L[D_{\pm}] = D'_{\pm}, \quad L[X] = X_1$$

$$(ii) \quad L A x = A_1 L x \quad x \in X, \quad A = P_X U|X, \quad A_1 = P_{X_1} U|X_1.$$

The result follows from the construction in Theorem 6.11.

CHAPTER SEVEN  
MARKOVIAN REPRESENTATIONS AND FACTORIZATION  
OF THE SCATTERING MATRIX

A regular Markov subspace  $X \subset H_{\underline{y}}$  is said to be a *representation* for the process  $\underline{y}$  if

$$\{y_1(0), \dots, y_p(0)\} \subset X.$$

Indeed, in this case we can write (notation as in Theorem 6.11)

$$\underline{y}(0) = [C]\underline{x}(0)$$

for some matrix  $[C]$ , and applying  $U$  to both sides gives

$$(7.1) \quad \underline{y}(n) = [C]\underline{x}(n),$$

a dynamical representation for the regular process  $\underline{y}$  in terms of the regular (and necessarily full range) Markov process  $\underline{x}$ . We first show that equivalent Markov processes are indistinguishable provided we do not distinguish between processes  $\underline{y}$ ,  $\underline{y}_1$  having the same spectral density (and see them as *equivalent*).

**7.1 Theorem.** Let  $X \subset H_{\underline{y}}$  be a regular Markovian representation for  $\underline{y}$ . If  $\{U^n X_1\}_{-\infty}^{\infty}$  is equivalent to  $\{U^n X\}_{-\infty}^{\infty}$ , there exists a process  $\underline{y}_1$  equivalent to  $\underline{y}$  for which  $X_1$  is a (regular) Markovian representation, and for which



(7.1) holds (with  $y_1, x_1$  playing the role of  $y, x$ ).<sup>+</sup>

Proof. Let  $L : H_y \rightarrow H_{y_1}$  be as in (6.11). By the properties of the incoming spectral representation

$$LU = UL \quad .$$

Define (componentwise)

$$y_1(0) = Ly(0) \quad , \quad y_1(n) = U^n y_1(0) \quad .$$

We obtain

$$\langle y_1(n), y_1(0) \rangle = \langle U^n Ly(0), Ly(0) \rangle = \langle LU^n y(0), Ly(0) \rangle = \langle y(n), y(0) \rangle \quad ,$$

and therefore

$$f_{y_1 y_1}(\lambda) = f_{yy}(\lambda) \quad \text{a.e. } \lambda \quad .$$

Since  $L[X] = X_1$ ,  $X_1$  is a regular Markovian representation for  $y_1$  and by Theorem 6.12 (7.1) holds. ◀

By the last theorem and Theorem 6.12 we can translate the question of finding all regular Markovian representations for  $y$  to  $L_2(C^P)$ . For each inner function  $\Theta \in H_\infty(B(C^P))$  (as in Theorem 6.10) we consider the orthogonal decomposition

$$L_2(C^P) = \Theta^* H_2^-(C^P) \oplus \left( (H_2^-(C^P) \ominus \Theta^* H_2^-(C^P)) \oplus H_2(C^P) \right)$$

---

<sup>+</sup> Note that the converse is trivial, see Theorem 6.12.

inducing the L-P scattering system

$$(\lambda, L_2(C^P), H_2(C^P), \ominus^* H_2^-(C^P)) ,$$

corresponding to the full range regular Markov subspace<sup>+</sup> (with respect to  $\lambda$ )

$$X = H_2^-(C^P) \ominus \ominus^* H_2^-(C^P) .$$

Now recall that for the incoming spectral representation  $F_-$  (for the scattering system associated with the  $y$  process)

$$F_- y(0) = \bar{\lambda} \Lambda .$$

Thus, finding all regular Markovian representations for  $y$  reduces to the following:

Find all inner functions  $\Theta_1 \in H_\infty(B(C^P))$  such that<sup>++</sup>

$$(7.2) \quad \bar{\lambda} \Lambda \in H_2^-(C^P) \ominus \ominus_1^* H_2^-(C^P) .$$

**7.2 Theorem.** All regular Markovian representations of  $y$  are parametrized by those and only those inner functions  $\Theta_1$  for which

---

<sup>+</sup> The corresponding decomposition in which  $X = H_2(C^P) \ominus \ominus H_2(C^P)$  (the model space, cf. [37]) will not be used here since we wish  $H_S$  to play a central role.

<sup>++</sup> Since  $\Lambda$  is conjugate outer the corresponding Markov subspace will automatically be of full range.

$$(7.3) \quad S = \Theta_1^* \Theta_2 \quad \Theta_2 \in H_\infty(B(C^p)) \quad . \quad ^+$$

Proof. (7.3) holds iff  $\Theta_1 \bar{\lambda} \in H_2(C^p)$  iff  $\Theta_1 S \Gamma \in H_2(C^p)$ . Since  $\Gamma$  is outer the latter holds iff  $\Theta_1 S \in H_2(C^p)$ . Since  $\Theta_1 S \in L_\infty(B(C^p))$  the result follows.  $\blacktriangleleft$

The possibility of writing the scattering matrix  $S$  in the form (7.3) has an interpretation on a process level. By the Beurling-Lax theorem, (7.3) holds iff (the invariant subspace for the left shift):

$$(7.4) \quad H_2^-(C^p) \ominus (\text{range } H_S) \text{ is of full range (for } \lambda).$$

By Lemma 4.2 this is equivalent to  $^{++}$

$$(7.5) \quad H_{\underline{y}}^-(0) \ominus P_{H_{\underline{y}}^-} H_{\underline{y}}^+(0) \text{ is of full range (for } U).$$

An  $L_\infty(B(C^p))$  function satisfying (7.4) is called *strictly non-cyclic* ([12], [15]). The process  $\underline{y}$  satisfying (7.5) (i.e., having a strictly non-cyclic scattering matrix) is called *strictly non-cyclic* [29].

Recall that two inner functions  $U, V \in H_\infty(B(C^p))$  are left coprime if for no non-trivial [i.e., for no unitary matrix in  $B(C^p)$ ] inner function  $W$  we have

$$U = WU_1, \quad V = WV_1$$

with  $U_1, V_1$  inner. Let

$^+$   $\Theta_2$  is necessarily inner since  $S$  is unitary valued on  $T$ .

$^{++}$  For the  $p = 1$  dimensional case the equivalence of (7.4) and (7.5) is due to Levinson-McKean [28].

$$S = Q_1^* Q_2$$

be the (left) coprime factorization for  $S$ . For the inner function  $Q_1$  we have by the Beurling-Lax theorem

$$\overline{\text{range } H_S} = H_2^-(C^P) \ominus Q_1^* H_2^-(C^P) .$$

If  $S = U_1^* U_2$  is some other factorization, then we necessarily have for some (non-trivial) inner function  $W$

$$U_1 = WQ_1 , \quad U_2 = WQ_2 .$$

Now observe that for inner functions  $U_1, Q_1$  in  $H_\infty(B(C^P))$

$$U_1^* H_2^-(C^P) \subset Q_1^* H_2^-(C^P)$$

iff

$$Q_1 U_1^* H_2^-(C^P) \subset H_2^-(C^P) ,$$

i.e., iff for some inner function  $W$

$$(7.6) \quad U_1 = WQ_1 .$$

It follows that for the (regular full range) Markov subspaces

$$X_{Q_1} = H_2^-(C^P) \ominus Q_1^* H_2^-(C^P), \quad X_{U_1} = H_2^-(C^P) \ominus U_1^* H_2^-(C^P)$$

we have

$$X_{Q_1} \subset X_{U_1}$$

Thus,  $X_{Q_1}$  is the *smallest* (setwise) regular Markov subspace representing  $\underline{y}$ . Let us call a regular Markov subspace  $X \subset H_{\underline{y}}$  representing  $\underline{y}$  *minimal* if no proper subspace of  $X^+$  is a Markovian representation of  $\underline{y}$ . Note that for the minimal subspace  $X_{Q_1}$  we have the orthogonal decomposition

$$L_2(C^P) = Q_1^* H_2^-(C^P) \oplus (H_2^-(C^P) \ominus Q_1^* H_2^-(C^P)) \oplus H_2(C^P)$$

which under  $F_-^{-1}$  goes to

$$(7.7) \quad H_{\underline{y}} = \left( H_{\underline{y}}^-(0) \ominus \overline{P_{H_{\underline{y}}^+(0)}} \right) \oplus \overline{P_{H_{\underline{y}}^-(0)} H_{\underline{y}}^+(0)} \oplus H_{\underline{y}}^+(0) \ominus H_{\underline{y}}^-(0) .$$

We established the following:

7.3 Theorem. Let  $\underline{y}$  be strictly non-cyclic, and

$$S = Q_1^* Q_2$$

the (left) coprime factorization of  $S$ . Then the equivalence class of all regular Markovian representations of  $\underline{y}$  (or an equivalent process to it) which corresponds to  $Q_1$  is minimal. This class is represented in  $L_2(C^P)$  by the subspace

$$X_{Q_1} = H_2^-(C^P) \ominus Q_1^* H_2^-(C^P) = \overline{\text{range } H_S} .$$

---

<sup>+</sup> Which is necessarily regular.

For each such Markov subspace, there exists a process  $\underline{\xi}$  equivalent to  $\underline{y}$  for which (7.1) holds. <sup>+</sup>

In a similar fashion, one obtains that all factorizations

$$S = U_2 U_1^*$$

parametrize via the inner function  $U_1$  all regular Markov subspaces  $X \subset H_{\underline{y}}$  representing  $\underline{y}$ . The minimal one is obtained by the (right) factorization  $S = P_2 P_1^*$  and its representation in  $L_2(C^P)$  is  $X_{P_1} = H_2(C^P) \ominus P_1 H_2(C^P)$  (see footnote on page 60).

A number of observations are in order. First, note that the general degree theory for strictly non-cyclic functions [15, Ch. III.5] arises naturally in our context. All regular Markovian subspaces  $X \subset H_{\underline{y}}$  representing  $\underline{y}$  which are in the equivalent class parametrized by the (purely contractive) inner function  $Q \in H_{\infty}(B(C^P))$  will be of degree

$$(7.8) \quad d(Q) = \det Q, \quad ,$$

an inner function in  $H_{\infty}$  (its scalar multiple, see Nagy-Foiaş [30, p. 216]).

From (7.6) we obtain

$$d(U_1) = d(Q_1) d(W), \quad ,$$

and thus the degree of the minimal subspace is the lowest in the sense that  $d(Q)$  is the weakest (i.e., it is an inner divisor) among the degrees of all other regular Markovian subspaces representing  $\underline{y}$ . Second, we

---

<sup>+</sup> With  $\xi$  playing the role of  $\underline{y}$ , and  $H_{\underline{\xi}} = H_{\underline{y}}$ .

observe that by inspection [see (7.7)], all minimal Markovian subspaces representing  $\underline{y}$  are observable and constructible<sup>+</sup> and conversely, all observable and constructible regular Markov subspaces representing  $\underline{y}$  are minimal. Moreover, we have exact observability and constructibility iff range  $H_S$  is closed.

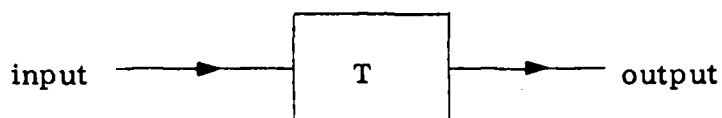
---

<sup>+</sup> The regular Markov subspace  $X$  is said to be *observable* if  $P_X H_Y^+(0) = X$ , and *constructible* if  $P_X H_Y^-(0) = X$ . It is said to be *exactly observable* or *constructible* if  $P_X H_Y^+(0) = X$ ,  $P_X H_Y^-(0) = X$  respectively. See [37].

## CHAPTER EIGHT

### MODEL REDUCTION

The question of model reduction for stationary processes is distinct in nature from the corresponding one in systems theory. In the latter, model reduction has a natural formulation in terms of inputs and outputs. Thus, the system is represented by a black box:



whose output  $\psi \in l_2(-\infty, \infty; C^q)$  when inputed by  $u \in l_2(-\infty, \infty; C^p)$  is given by

$$\psi = Lu \quad ,$$

where  $L : l_2(-\infty, \infty; C^p) \rightarrow l_2(-\infty, \infty; C^q)$  is a bounded causal  $[Ll_2(0, \infty; C^p) \subset l_2(0, \infty; C^q)]$  Laurent operator (input-output map) with symbol  $T \in H_\infty(B(C^p, C^q))$  (frequency response function). The (minimal) state space of the system, i.e., the lowest dimension of  $\underline{x}$  in the dynamical representation <sup>+</sup>

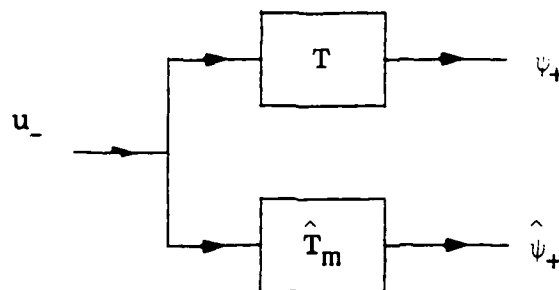
$$\begin{aligned} \underline{x}(n+1) &= [A]\underline{x}(n) + [B]u(n) \\ (8.1) \quad \psi(n) &= [C]\underline{x}(n) \end{aligned} \quad ,$$

---

<sup>+</sup> We assume  $T(0) = [0]$ .



equals  $\text{rank } H_T = k$ .<sup>+</sup> One way to formulate a model reduction problem in this context is the following: Given  $T$  find a (system) function  $\hat{T}_m \in H_\infty(B(C^P, C^Q))$  (with corresponding Laurent operator  $\hat{L}_m$ ) with  $\text{rank } H_{\hat{T}_m} \leq m < k$ , such that if the two systems were input by the same  $u_- \in l_2(-\infty, -1; C^P)$  up to time  $t = 0$ :



for the corresponding outputs from time  $t = 0$  onward :  $\psi_+, \hat{\psi}_+ \in l_2(0, \infty; C^Q)$ , we have

$$\|\psi_+ - \hat{\psi}_+\|_{l_2} = \min.$$

over all  $u_- \in l_2(-\infty, -1; C^P)$ ,  $\|u_-\| \leq 1$  (normalization condition). Taking Fourier transforms, we readily obtain

$$\min_{\hat{T}_m} \sup \{ \|(H_T - H_{\hat{T}_m})\phi\|_{L_2} : \phi \in H_2^-(C^P), \|\phi\|_{L_2} \leq 1 \}$$

where the min. is taken over all  $\hat{T}_m \in H_\infty(B(C^P, C^Q))$  such that  $\text{rank } H_{\hat{T}_m} \leq m$ . Thus, we have

---

<sup>+</sup>  $H_T$  being the Hankel operator from  $H_2^-(C^P) \rightarrow H_2(C^Q)$ .

$$\min_{\hat{T}_m} \|H_T - \hat{H}_{T_m}\| ,$$

a classical Hankel approximation problem [7].

This input-output notion is not adequate in the context of stationary sequences. While a realization

$$\begin{aligned} \underline{x}(n+1) &= [A]\underline{x}(n) + [B]\underline{w}(n) \\ (8.2) \qquad \qquad \qquad n &\geq 0 \\ \underline{y}(n) &= [C]\underline{x}(n) \end{aligned}$$

resembles the one in systems theory, one should keep in mind that the notion of inputs is absent, and the sequence  $\{\underline{y}(n)\}_{-\infty}^{\infty}$  (or a sequence equivalent to it) plays a distinct role. One approach close in nature to the above model reduction formulation is to approximate this sequence. We wish, however, to emphasize that while we formulate the approximation problem geometrically in the space  $H_{\underline{y}}$ , the object we seek is the approximating spectral density. Thus, all approximating sequences which follow will 'live' in  $H_{\underline{y}}$  and will be, with respect to the unitary operator  $U$ , on  $H_{\underline{y}}$ . The resulting processes will be, therefore, stationary correlated with  $\underline{y}$ . Observe that approximating the sequence  $\{\underline{y}(n)\}_{-\infty}^{\infty}$  or an equivalent sequence to it are thus indistinguishable.

**Problem A.** Let  $\{\underline{y}(n)\}_{-\infty}^{\infty}$  be the given process with (minimal) representation (8.2),  $\dim \underline{x} = \text{rank } H_S = k \leq \infty$ . Find a regular process  $\{\hat{\underline{y}}(n)\}$  of dimension  $p$  such that for its corresponding (minimal) representation

$$\begin{aligned}\hat{\underline{x}}(n+1) &= [\hat{A}]\hat{\underline{x}}(n) + [\hat{B}]\underline{w}(n) \\ \hat{\underline{y}}(n) &= [\hat{C}]\hat{\underline{x}}(n)\end{aligned}\quad n \in \mathbb{Z},$$

we have

$$(i) \quad \dim \hat{\underline{x}} \leq m < k,$$

and, moreover

$$(ii) \quad \sup \{ \|\underline{y}_j(n) - \hat{\underline{y}}_j(n)\|_{H_{\underline{y}}} : j=1, \dots, p, n \in \mathbb{Z} \} = \min.$$

Assume first  $p=1$ . Let  $\{\hat{\underline{y}}(n)\}_{-\infty}^{\infty}$  be a solution and  $\hat{X}$  one of its minimal regular Markovian representations ( $\dim \hat{X} \leq m$ ). Define

$$\tilde{\underline{y}}(0) = P_{\hat{X}} \underline{y}(0).$$

By the orthogonality principle, we obtain

$$\hat{\varepsilon}^2(0) = \|\underline{y}(0) - \hat{\underline{y}}(0)\|^2 \geq \|\underline{y}(0) - \tilde{\underline{y}}(0)\|^2 = \tilde{\varepsilon}(0)^2$$

Clearly for the process

$$\{U^n \tilde{\underline{y}}(0) : n \in \mathbb{Z}\}$$

the (minimal) state space is of dimension not exceeding  $m$ . By stationarity

$$\tilde{\varepsilon}(n) = \|\underline{y}(n) - \tilde{\underline{y}}(n)\| = \|U^n \underline{y}(0) - U^n \tilde{\underline{y}}(0)\| = \tilde{\varepsilon}(0)$$

and

$$\hat{\varepsilon}(0) = \sup \{ \|\underline{y}(n) - \hat{\underline{y}}(n)\|^2 : n \in \mathbb{Z} \} \geq \tilde{\varepsilon}(0).$$

We readily conclude

$$\hat{y}(0) = P_{\hat{X}} y(0) \quad .$$

As for arbitrary  $p$ , we note that the above procedure extends componentwise for each  $y_j$   $j=1, \dots, p$ , and thus for the process  $\underline{\hat{y}}$  we need to have

$$\hat{y}_j(0) = P_{\hat{X}} y_j(0) \quad j=1, \dots, p \quad .$$

Our problem thus reduces (in the space  $H_{\underline{y}}$ ) to the following:

Find a regular Markov subspace  $\hat{X} \subset H_{\underline{y}}$  such that:

- (8.3) (i)  $\dim \hat{X} \leq m$   
 (ii)  $\max_{j=1, \dots, p} \text{dist}_{H_{\underline{y}}} (y_j(0), \hat{X}) = \min.$

Observe that if  $\epsilon$  is the minimum above

$$\begin{aligned} |(C_n)_{i,j} - (\hat{C}_n)_{i,j}| &= |(y_i(n), y_j(0)) - (\hat{y}_i(n), \hat{y}_j(0))| \leq \\ &\leq \|(y_i(n), y_j(0) - \hat{y}_j(0))\| + \|(y_i(n) - \hat{y}_i(n), \hat{y}_j(0))\| \leq 2\epsilon \sigma_{y_i y_i} \end{aligned}$$

Thus for all  $n$

$$\|C_n - \hat{C}_n\| \leq \left[ (2\epsilon)^2 \sum_{i,j=1}^p \sigma_{y_i y_i}^2 \right]^{\frac{1}{2}} = 2\epsilon \sqrt{p} \left[ \sum_{i=1}^p \sigma_{y_i y_i} \right]^{\frac{1}{2}} .$$

Since  $\hat{y}$  is regular and of full rank  $p$ , it follows from [30, Prop. 1.2.1] that  $\hat{X}$  is of full range and, therefore, belonging to the equivalent class of regular full range Markov subspaces parametrized by the (purely contractive) inner functions  $\Theta \in H_{\infty}(B(C^p))$  for which  $d(\Theta)$  [see

(7.8)] is a Blaschke product of degree  $\leq m$ . From Chapter 7, (8.3) translates in  $L_2(C^P)$  into finding a  $Q$  for which

$$\max_{j=1, \dots, p} \text{dist}_{L_2(C^P)}(\bar{\chi} \Lambda_j, H_2^-(C^P) \ominus Q^* H_2^-(C^P)) = \min., \quad \Lambda = [\Lambda_1 | \dots | \Lambda_p].$$

**8.1 Proposition.** Let  $\theta \in H_\infty(B(C^P))$  be an inner function and  $g \in H_2^-(C^P)$ .

Then

$$P_{H_2^-(C^P)} \ominus \theta^* H_2^-(C^P) g = \theta^* \pi_+ \theta g.$$

Proof. One readily checks that  $\theta^* \pi_+ \theta g \in H_2^-(C^P)$ . Since

$$g - \theta^* \pi_+ \theta g = \theta^* [\theta g - \pi_+ \theta g] = \theta^* [(I_{L_2} - \pi_+) \theta g] = \theta^* \pi_- \theta g \in H_2^-(C^P),$$

the result follows. ◀

Now observe that

$$\bar{\chi} \Lambda - \theta^* \pi_+ \theta \bar{\chi} \Lambda = \theta^* [\theta \bar{\chi} \Lambda - \pi_+ \theta \bar{\chi} \Lambda] = \theta^* \pi_- (\theta \bar{\chi} \Lambda) = \theta^* H_{\theta S}^\Gamma.$$

Problem A therefore admits the following equivalent formulation:

Problem A~. Find an inner function  $Q \in H_\infty(B(C^P))$  such that  $\|Q(0)\| < 1$ , and the following holds:

$$(8.4) \quad \begin{aligned} & \text{(i)} \quad d(Q) \leq m^+, \\ & \text{(ii)} \quad \max_{j=1, \dots, p} \|H_{QS}^{\Gamma_j}\|_{L_2} = \min. \quad \Gamma = [\Gamma_1 | \dots | \Gamma_p]. \end{aligned}$$

---

<sup>+</sup> i.e., the Blaschke product is of degree  $\leq m$ .

We first establish the existence of  $Q$  solving (8.4) for the case  $p = 1$ .

8.2. Theorem. There exists an inner function  $q^+$  of degree  $\leq m$  such that

$$(8.5) \quad \|H_{Q_S} \Gamma\| = \min \{ \|H_{u_S} \Gamma\| : u \text{ inner, degree } u \leq m \}.$$

Proof. Denote by  $\delta$  the infimum of the right hand side in (8.5) and let  $\{u_n\}_0^\infty$  be a sequence of Blaschke products such that  $\|H_{u_n S} \Gamma\| \rightarrow \delta$ . The family  $u_0, u_1, \dots$  is uniformly bounded in  $|z| < 1$ , in particular, on compact subsets of  $|z| < 1$ . It follows (e.g., [39, Th. 14.6]) that there exists a subsequence  $\{u_{n_j}\}_{j=0}^\infty$  converging uniformly on compact subsets of  $|z| < 1$  to a function  $q \in H_\infty$  which is  $\neq 0$ . By a theorem of Rouché [39, p. 242], it readily follows that if  $|a| < 1$  is a zero of  $q$ , then there exists a sequence  $\{a_{n_j}\}_{j=0}^\infty$  in  $|z| < 1$  such that  $\lim_{j \rightarrow \infty} a_{n_j} = a$  and for some  $N$

$$u_{n_j}(a_{n_j}) = 0 \quad j \geq N.$$

Since each  $u_n$  has at most  $m$  zeros in  $|z| < 1$ ,  $q$  has at most  $m$  zeros in  $|z| < 1$ . Let  $0 < r_0 < 1$  be such that  $q$  has no zeros in  $r_0 \leq |z| < 1$ . Applying once more Rouché's theorem, we conclude that for some integer  $N_0$ , all  $u_{n_j}$ ,  $j \geq N_0$  have the same number of zeros in  $r_0 \leq |z| < 1$  as  $q$ , i.e., none. This implies that the family  $\{u_{n_j}\}_{j \geq N_0}$  is holomorphic in a region containing  $|z| < 1$ . We readily conclude that  $q$  is an inner function having at most  $m$  zeros in  $|z| < 1$  and none on  $|z| = 1$ , i.e., a Blaschke product of degree  $\leq m$ .

---

<sup>+</sup> Clearly  $q$  is non-trivial.

As for general  $p$ , by a normal family argument one obtains a subsequence  $\Theta_{n_j}$  converging uniformly over compact subsets of  $|z| < 1$  [uniform topology in  $B(C^p)$ ] to an inner function  $Q$ . Applying the argument of the above proof to the inner functions  $d(\Theta_{n_j})$  which converge (uniformly over compact subsets of  $|z| < 1$ ) to the inner function  $d(Q)$ , we conclude that  $Q$  is of degree  $d(Q)$  (a Blaschke product of degree)  $\leq m$ . To show that it is purely contractive, we note that if for some  $0 \neq a \in C^p$   $\|Q(0)a\| = \|a\|$  then  $Q(z)a \in H_2(C^p)$  is a constant function and thus, considering  $\|Q(z)a\|_{\epsilon H_2}$  we obtain  $\|\Theta_{n_j}(z)a\| \rightarrow \|Q(z)a\|$  uniformly over compact subsets. By Rouché's theorem, this is impossible since the  $\Theta_{n_j}$  are purely contractive. We conclude that  $Q$  is purely contractive.

The solution obtained is not necessarily unique, and, moreover, it is not obtained in a constructive fashion. A weaker version (which is nonetheless no less interesting) is obtained by trying to minimize  $\|H_{\Theta_S}\|$  over all admissible  $\Theta$ . For the  $p = 1$  dimensional case, a constructive solution can be obtained by employing the analytic properties of Schmidt pairs for a Hankel operator [7]. Thus, since

$$\|H_{us}\Gamma\| \leq \|H_{us}\| \|\Gamma\|$$

we have

$$\min\{\|H_{us}\Gamma\| : u \text{ inner, degree } u \leq m\} \leq \min\{\|H_{us}\| : u \text{ inner, degree } u \leq m\} \cdot \|\Gamma\|$$

and we consider

**Problem B.** Find an inner function  $^+ q$  such that

---

<sup>+</sup> This inner function  $q$  will be found to be non-trivial (see Theorem 6.10).

$$(8.6) \quad \begin{aligned} & \text{(i)} \quad \text{degree } q \leq m \\ & \text{(ii)} \quad \|H_{qS}\| = \min . \end{aligned}$$

We make the assumption (which we later characterize on a process level) that  $H_S$  is a compact operator. Let

$$1 = s_0(H_S) \geq s_1(H_S) \geq \dots$$

be the enumeration of its non-zero singular numbers (s-number) repeated with multiplicity (which is finite by the compactness). These are the positive square root of the eigenvalues of

$$[H_S]^* [H_S] .$$

If  $H_S$  is of finite rank  $k < \infty$ , 0 is an s-number with infinite multiplicity and the enumeration will be

$$1 = s_0(H_S) \geq \dots \geq s_{k-1}(H_S) > s_k(H_S) = s_{k+1}(H_S) = \dots = 0.$$

8.3 Theorem. Let

$$\dots \geq s_{m-1}(H) > s_m(H) \geq \dots ,$$

then there exists a unique  $^+ q$  of exact degree  $m$  satisfying (8.6).

Proof. Denote by  $B_m$  the family of all Blaschke products of degree  $\leq m$ .

By a result of Kronecker (cf. [35]) for  $\phi \in L_\infty$  rank  $H_\phi \leq m$  iff  $\phi = \bar{u}h$

---

$^+$  Up to coincidence.



with  $u \in B_m$ ,  $h \in H_\infty$ . Combining with Nehari's theorem, we obtain

$$\begin{aligned} \inf_{u \in B_m} \|H_{us}\| &= \inf_{u \in B_m} \inf_{h \in H_\infty} \|us - h\|_{L_\infty} = \inf_{u \in B_m} \inf_{g, h \in H_\infty} \|us - h - ug\|_{L_\infty} = \\ &= \inf_{g \in H_\infty} \inf_{\substack{u \in B_m \\ h \in H_\infty}} \|(s - \bar{u}h) - g\|_{L_\infty} = \inf\{\|H_s - H_\psi\| : \text{rank } H_\psi \leq m\} = s_m(H_s). \end{aligned}$$

Now assume that  $u_1, u_2$  satisfy (8.6). By Nehari's theorem we obtain for some  $h_1, h_2 \in H_\infty$

$$s_m(H_s) = \|H_{u_j s}\| = \|u_j s - h_{u_j}\|_{L_\infty} = \|s - \bar{u}_j h_j\|_{L_\infty} \quad j=1,2.$$

It follows that for the Hankel operators  $H_{\bar{u}_j h_j}$  (of rank  $\leq m$ ) we have

$$\|H_s - H_{\bar{u}_j h_j}\| \leq \|s - \bar{u}_j h_j\|_{L_\infty} = s_m(H_s) \quad j = 1, 2,$$

and, therefore,

$$s_m(H_s) = \|H_s - H_{\bar{u}_1 h_1}\| = \|H_s - H_{\bar{u}_2 h_2}\|.$$

By a theorem of Adamjan-Arov-Krein [7, Th. 1.1] there exists a unique Hankel operator  $H_{\phi_m}$  of rank  $m$  such that  $\|H_s - H_{\phi_m}\| = s_m(H_s)$ . We conclude that

$$H_{\phi_m} = H_{\bar{u}_1 h_1} = H_{\bar{u}_2 h_2}.$$

Let  $q$  be the inner function of degree  $m$  such that

$$\text{range } H_{\phi_m} = H_2^- \ominus \bar{q} H_2^- .$$

It follows that  $q$  is an inner divisor of  $u_1$  and  $u_2$ . Since  $u_1, u_2$  are of degree  $\leq m$  we conclude that  $u_1, u_2, q$  coincide which completes the proof. ◀

**8.4 Remark.** Note that for  $m = k = \text{rank } H_s$  the inner function  $q$  coincides with the inner function  $q_1$  in the coprime factorization  $s = \bar{q}_1 q_2$  (see Theorem 7.3).

This  $q$  is induced by the unique rank  $m$  Hankel approximant  $H_{\phi_m}$  via

$$X_q = H_2^- \ominus q H_2^- = \text{range } H_{\phi_m} .$$

On this Markov subspace in  $L_2$  (with respect to  $\chi$ ), we project  $F_- y(0) = \bar{x} \bar{\Gamma}$ , and thereby obtain the approximating process  $y$  induced by  $q$ . Indeed, in the space  $H_y$  we have

$$H_y \supset X_q = F_-^{-1}[X_q] .$$

$y(0)$  is obtained by

$$\hat{y}(0) = P_{X_q} y(0)$$

and we set  $\hat{y}(n) = U^n \hat{y}(0)$  .

Incorporating [7], we thus have the following procedure of constructing the approximating process  $\hat{y}$  with its density  $\hat{f}_{\hat{y}\hat{y}}$  which are

induced by  $q$ :

Step 1. Let  $s_{m-1}(H_s) > s_m(H_s)$  have multiplicity  $\mu$ , and  $\xi_1, \xi_2, \dots, \xi_\mu \in l_2(0, \infty; C)$  be the corresponding eigenvectors, i.e.,

$$[H_s]^*[H_s]\xi_j = s_m^2(H_s)\xi_j \quad j = 1, 2, \dots, \mu.$$

Let

$$f_j(z) = \sum_{n=0}^{\infty} \xi_j(n) z^n \epsilon_{H_2} \quad j = 1, 2, \dots, \mu.$$

Factor

$$f_j(z) = u_j(z) \psi_j(z)$$

where  $u_j$  is inner,  $\psi_j$  outer. The desired function  $q$  is

$$q = u_1 \wedge u_2 \wedge \dots \wedge u_\mu,$$

i.e., the greatest common inner divisor for the  $u_j$  [7, Th. 1.2]. By Proposition 8.1 the projection of  $\bar{\chi}\bar{\Gamma}$  onto  $X_q$  is given by

$$\bar{q} \pi_+ q \bar{\chi} \bar{\Gamma}.$$

Step 2. Factor

$$\bar{q} \pi_+ q \bar{\chi} \bar{\Gamma} = \bar{\chi} \bar{\phi}_i \bar{\psi}_e$$

where  $\phi_i$  is inner,  $\psi_e$  outer. Observe that

$$\begin{aligned}\hat{c}_n &= (\hat{y}(n), \hat{y}(0))_{H_y} = (F_{-}\hat{y}(n), F_{-}\hat{y}(0))_{L_2} = (\chi^n \bar{\chi} \bar{\phi}_i \bar{\psi}_e, \bar{\chi} \bar{\phi}_i \bar{\psi}_e) = \\ &= (\chi^n \psi_e, \psi_e) \quad .\end{aligned}$$

Thus  $\psi_e$  is an outer factor for

$$f_{\hat{y}\hat{y}}(\lambda) = \frac{1}{2\pi} |\psi_e(e^{i\lambda})|^2 \quad ,$$

the density of the  $\hat{y}$  process.

We next obtain a bound on the normalized <sup>+</sup> difference

$$\left| \frac{c_n - \hat{c}_n}{c_0} \right|$$

between the moments  $c_n = (y(n), y(0))$ ,  $\hat{c}_n = (\hat{y}(n), \hat{y}(0))$  of the corresponding  $y$ ,  $\hat{y}$  processes.

**8.5 Proposition.** We have

$$(8.7) \quad \left| \frac{c_n - \hat{c}_n}{c_0} \right| \leq 2s_m(H_s) \quad , \quad n = 0, \pm 1, \dots \quad .$$

**Proof.** From the discussion preceding problem A~ and the proof of Theorem 8.3, we obtain

$$\begin{aligned}\|y(n) - \hat{y}(n)\|_{H_y} &= \|\hat{y}(0) - \hat{y}(0)\| = \|H_{qs}\Gamma\|_{L_2} \leq \\ &\leq \|H_{qs}\| \|\Gamma\| = s_m(H_s) \sqrt{c_0} \quad .\end{aligned}$$

---

<sup>+</sup> To the variance  $\sigma_{yy}^2 = c_0$  of the  $y$  process.

Thus,

$$\begin{aligned}
 \left| \frac{c_n - \hat{c}_n}{c_0} \right| &= \left| \frac{(y(n), y(0)) - (\hat{y}(n), \hat{y}(0))}{c_0} \right| \leq \\
 &\leq \frac{1}{c_0} \left[ |(y(n), y(0) - \hat{y}(0))| + |(\hat{y}(n) - y(n), \hat{y}(0))| \right] \leq \\
 &\leq s_m(H_s) + s_m(H_s) \frac{\sqrt{c_0} \sqrt{\hat{c}_0}}{c_0} \leq 2s_m(H_s)
 \end{aligned}$$

We make a number of observations. First, observe that

$$\|y(0) - \hat{y}(0)\|_{H_y} = \|\Gamma - \phi_i \psi_e\|_{L_2}.$$

Thus,  $\psi_e$  can not be seen as a rational approximation in  $L_2$  norm to the outer factor  $\Gamma$  of  $f_{yy}$ . If rank  $H_s$  is infinite or otherwise if the singular values drop sharply, (8.7) demonstrates that the process  $\hat{y}$  is a rather good approximate to  $y$ . Second, we note that the assumption about the compactness of  $H_s$  is not a necessary one. Indeed, the spectrum of a bounded Hankel operator is the union of its point spectrum and its essential spectrum [34]. The analytic properties of Schmidt pairs were given in [7] for a general bounded Hankel operator and, thus, our construction carries over to this case *mutatis-mutandis*. The multiplicity  $\mu$  of  $s_m(H_s)$ , however, may be infinite in this case.

We now characterize the compactness condition on  $H_s$  on a process level. For any two subspaces  $M, N$  of a Hilbert space  $H$  let

$$\rho(M, N) = \sup \{ |(f, g)_H| : f \in M, g \in N, \|f\|, \|g\| \leq 1 \}.$$

It readily follows that  $\rho(M, N) = \|P_M P_N\|$ . If  $\rho < 1$ ,  $M, N$  are said to be

at positive angle [22]. A process  $y$  is said to be *strongly mixing* (Yaglom [43]) if

$$\rho(H_y^-(0), H_y^+(n)) \rightarrow 0 \quad (n \rightarrow \infty),$$

i.e., its distant past and future are nearly orthogonal. It is known (see e.g. [22]) that a regular process is strongly mixing iff the operator  $P_- P_+$  is compact.

Combining with Lemma 4.2 we conclude by using [32]

**8.6 Proposition.**  $y$  is strongly mixing iff  $H_S$  is compact, i.e.,

$$S \in H_\infty(B(C^P)) + C(B(C^P)).^+$$

Processes with rational density are thus strongly mixing. Helson-Sarason [22] characterized the spectral density of (1 dimensional) processes which are strongly mixing

$$f_{yy} = |P|^2 \exp(u + \tilde{v})$$

where  $u, v$  are continuous on  $T$ ,  $\tilde{v}$  the Harmonic conjugate of  $v$ , and  $P$  is a polynomial with roots on  $T$ . In particular, densities which are continuous and strictly positive are in this class.

**8.7 Remark.** As for the vector generalization of the above construction, we remark that if  $H_{\phi_m}$  is the best rank  $m$  Hankel approximant to  $H_S$ ,

---

<sup>+</sup>  $C(B(C^P))$  denotes the continuous,  $B(C^P)$  valued functions  $f$  on the circle.

the solution can be carried over provided we guarantee that  $\phi_m$  is strictly noncyclic, i.e.,

$$\phi_m = \Theta^* P$$

$\Theta \in H_\infty(B(C^p))$  inner,  $P \in H_\infty(B(C^p))$ , and, moreover,  $\|\Theta(0)\| < 1$ . We did not find a way to guarantee it.

## REFERENCES

- [1] V.M. Adamjan. Non degenerate unitary couplings of semiunitary operators, Funct. Anal. Appl., 7 (1973), 255-267.
- [2] V.M. Adamjan and D.Z. Arov. On unitary couplings of semi-unitary operators, Amer. Math. Soc. Transl. 95(2) (1970), 17-129.
- [3] \_\_\_\_\_. A general solution of a problem in linear prediction of stationary processes, Theory of Prob. and its Appl. 13(3) (1968), 394-407.
- [4] V.M. Adamjan, D.Z. Arov, and M.G. Krein<sup>v</sup>. Infinite Hankel block matrices and related extension problems. Amer. Math. Soc. Transl. 3(2) (1978), 133-56.
- [5] \_\_\_\_\_. Infinite Hankel matrices and generalized Carathéodory-Fejer and Riesz problems, Funktsional'nyy Analiz i Ego Prilozheniya 2(1) (1968), 1-18.
- [6] \_\_\_\_\_. Infinite Hankel matrices and generalized Carathéodory-Fejer and I. Schur problems, Funkt. Anal. i Ego Prilozhen. 7(4) (1973), 1-16.
- [7] \_\_\_\_\_. Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem, Mat. Sb. 86 (128) (1971), 34-75; Math. U.S.S.R. Sb. (1971), 31-73.
- [8] J.A. Ball. Invariant subspace representations, unitary interpolants and factorization indices. Operator Theory: Adv. and Appl. 12 (1984), 11-38.
- [9] P.B. Bloomfield, W.P. Jewell, and E. Hayashi. Characterizations of completely nondeterministic stochastic processes, Pacific J. of Math. 107 (1983), 307-17.
- [10] J.R. Butz. s-Numbers of Hankel matrices. J. Functional Analysis 15 (1974), 297-305.
- [11] R.G. Douglas. Banach Algebra Techniques in Operator Theory. Academic Press, New York, 1972.
- [12] R.G. Douglas, H.S. Shapiro and A.L. Shields. Cyclic vectors and invariant subspaces for the backward shift, Ann. Inst. Fourier (Grenoble) 20(1) (1971), 37-76.
- [13] C. Foiaş, and E. Frazho. A note on unitary dilation theory and state spaces, Acta Sci. Math. 45 (1983), 165-75.
- [14] Y. Fourès and I.E. Segal. Causality and analyticity, Trans. Am. Math. Soc. 78 (1955), 385-407.



- [15] P.A. Fuhrmann. Linear Systems and Operators in Hilbert Space. McGraw-Hill, New York, 1981.
- [16] I.C. Gohberg and M.G. Krein<sup>y</sup>. Introduction to the Theory of Linear Nonselfadjoint Operators. Transl. of Math. Monograph 18, Amer. Math. Soc., 1969.
- [17] U. Grenander and Gabor Szegö. Toeplitz Forms and Their Applications. University of California Press, Berkeley, 1958.
- [18] P.R. Halmos. A Hilbert Space Problem Book. Van Nostrand, New York, 1967.
- [19] \_\_\_\_\_. Shifts on Hilbert spaces, J. Reine Angew Math. 208 (1961), 102-12.
- [20] P. Hartman. On completely continuous Hankel matrices, Proc. Amer. Math. Soc. 9 (1958), 362-66.
- [21] H. Helson. Lectures on Invariant Subspaces. Academic Press, New York, 1964.
- [22] H. Helson and D. Sarason. Past and future, Math. Scand. 21 (1967), 5-16.
- [23] J.W. Helton. Discrete time systems, operator models and scattering theory, J. Funct. Anal. 16(1) (1974), 15-38.
- [24] K. Hoffman. Banach Spaces of Analytic Functions. Prentice-Hall, Englewood Cliffs, 1962.
- [25] I.A. Ibragimov.. On the spectrum of stationary Gaussian sequences satisfying the strong mixing condition: I. Necessary conditions. Theory of Probability and its Applications 10(1) (1965), 85-106.
- [26] E.A. Jonckheere, and J.W. Helton. Power spectrum reduction by optimal Hankel approximations of the phase of the outer spectral factor, Proceedings of the American Control Conference, 3 (June, 1984), 1352-358.
- [27] P.D. Lax and R.S. Phillips. Scattering Theory. Academic Press, New York, 1967.
- [28] N. Levinson and H.P. McKean, Jr. Weighted trigonometrical approximations on  $R^1$  with applications to the germ field of stationary Gaussian noise, Acta. Math., 112 (1964), 99-143.
- [29] A. Lindquist and G. Picci. State space models for Gaussian stochastic processes. In M. Hazewinkel and J.C. Willems (eds.), Stochastic Systems: The Mathematics of Filtering and Identification and Applications. Reidel, 1981, 169-204.
- [30] B. Sz-Nagy and C. Foiaş. Harmonic Analysis of Operators on Hilbert Space. Amsterdam, North-Holland, 1970.

- [31] Z. Nehari. On bounded bilinear forms, Annals of Mathematics 65(1) (1957), 153-62.
- [32] L.B. Page, Bounded and compact vectorial Hankel operators, Trans. Amer. Math. Soc. 150 (1970), 529-39.
- [33] V.V. Peller. Hankel operators of class  $O_p$  and their applications (rational approximation, Gaussian processes, the problem of majorizing operators), Math. U.S.S.R. Sbornik 41(4) (1982), 443-79.
- [34] S.C. Power. Hankel operators on Hilbert space. Bull. London Math. Soc. 12 (1980), 422-42.
- [35] \_\_\_\_\_. Hankel Operators on Hilbert Space. Pitman Advanced Publishing Program, Boston, 1982.
- [36] Y.A. Rozanov. Stationary Random Processes. Holden-Day, San Francisco, 1963.
- [37] G. Ruckebusch. Théorie géométrique de la représentation Markovienne, Ann. Inst. Henri Poincaré, 16(3) (1980), 225-97.
- [38] \_\_\_\_\_. On the structure of minimal Markovian representations, In Nonlinear Stochastic Problems, R. Bucy and J.M.F. Moura (eds.), Reidel, 1983.
- [39] W. Rudin. Real and Complex Analysis. 2nd ed. McGraw Hill, New York, 1974.
- [40] D.E. Sarason. Function theory on the unit circle. Notes for Lectures at a conference at Virginia Polytechnic and State University, Virginia (1978).
- [41] \_\_\_\_\_. Generalized interpolation in  $H^\infty$ , Trans. Amer. Math. Soc. 127 (1967), 179-203.
- [42] J.C. Willems. System theoretic models for the analysis of physical systems. Ricerca di Automatica 10(2) (1979), 71-106.
- [43] A.M. Yaglom. Stationary Gaussian processes satisfying the strong mixing condition and best predictable functionals, Proc. Int. Research Seminar of Statistical Lab., Univ. of Calif., Berkeley, 1963, 241-52; Springer-Verlag, 1965.

**END**

**FILMED**

**11-85**

**DTIC**